Development of a Web-based Diagnosis System Using Bayesian Classifier Technique

Akinwole Agnes K.[1], Adetoba Bolaji T.[2], Ogundele Israel I.[3], Alakiri Harrison O.[4]
Yaba College of Technology
[1,2,3,4]Department of Computer Technology
Yaba, Lagos, Nigeria
E-mail: agnes.akinwole@gmail.com[1]
Phone: 234-803-713-2123

ABSTRACT

A web-based medical diagnosis system is a system that determines which diseases or condition explain a person's symptoms at anytime, anywhere by any person. Diagnosing a patient is a major challenge in the healthcare service delivery because a symptom might represent several diseases or disorder which do not give a precise information about what is wrong with the patient. This paper uses a Bayesian Classifier Technique to diagnose patient based on Twenty-Six dataset features which were used to predict the likelihood of any of the five common diseases (Malaria, Typhoid, Yellow Fever, Cholera and Tuberculosis) among students in Yaba College of Technology. Visual Analogue Scale for medical diagnosis was used to evaluate the result of the Bayesian Classifier in no disease, mild disease, moderate and severe disease for the severity level of the diseases which determines the recommendation for further treatment.

Keywords: Bayesian Classifier, Diseases, Medical diagnosis, Web-based, Visual Analogue Scale.

1. BACKGROUND OF THE STUDY

The development and application of Information Communication Technology (ICT) in the health sector of developing countries has brought about a tremendous change in the total well-being of an individual. Health is very important for every human being since it shows the state of healthy living of a person. Unfortunately, almost everyone nowadays is suffering from at least one type of illness, particularly diseases that are common to adolescents or young adults such as Malaria, Typhoid, Yellow Fever, Cholera and Tuberculosis. It is a challenge situation for a government in developing nation to find medical experts to handle enormous patients that need medical attention daily. A Web Based Medical Diagnosis System is a system that uses web interface and expert technology to act as a human expertise in order to diagnose possible diseases based on given symptoms (Arbaiy, Sulaiman, Hassan & Afip, 2017).

Medical expert systems are developed with several application systems to complement the duties of health workers and practitioners. These systems have been applied in a wide range of industrial and commercial problems including diagnosis, planning, scheduling, decision support, process monitoring, classification and control. In medical science, Bayesian Classifier can be used as the logical process of performing medical diagnosis, particularly in automated medical diagnosis decision support systems (De Silva & Jayamanne, 2016). It describes the probability of an event based on prior knowledge of conditions that might be related to the event. A Visual Analogue Scale (VAS) is a measuring instrument used to measure characteristic or attitude that is believed to range across a scale of values and cannot easily be directly measured.

It is often used in epidemiologic and clinical research to measure the strength or occurrence of various symptoms (Streiner, Norman & Cairney, 2015. This paper reviewed previous works on medical diagnosis system and developed a Web-Based Medical Diagnosis System using Bayesian Classifier Techniques. The system gives the diagnosis as the doctor gives to the patient with or without the presence of the physician at any time. The System automates the whole process of patient's diagnosis with accuracy and present the severity using Bayesian Classifier to achieve desired output. The VAS was used to evaluate the result of the Bayesian classifier into their level of severity (no disease, mild disease, moderate and severe disease) and the result is used to determine further recommendation for the patient. This will prevent unnecessary visit of the patient to the medical Centre thereby reduce the work stress of the health workers.

1.1 Statement of the Problem

It is a challenging condition in YCT on how students fall ill during academic session especially during examination period. This result to increase in students admission at the college medical Centre which give a lot of stress to the medical staffs and physicians. Due to the stress faced by the health workers, wrong diagnosis are sometimes prevalent which may at times result to incorrect prescription. When a patient is given an incorrect prescription, it may lead to serious health issues and if proper care is not taken in due time, this may result to death. This paper therefore focused on developing a web-based medical diagnosis system to assist both the students and medical Centre health worker in Yaba College of Technology in order to get a real time medical health status.

1.2 Aim and Objectives

The aim of this paper is to develop a web-based medical diagnosis system using Bayesian classifier Technique.

Objectives include

i) review previous work on medical diagnosis system in order to improve on the existing framework
ii) develop a web-based medical diagnosis system using Bayesian classifier
iii) using Visual Analogue Scale (VAS) to evaluate the severity of the diseases for further recommendation

2. RELATED WORKS

Various researchers have acknowledged the fact that there is a need for medical diagnosis system to assist the health workers and physicians in discharging their duties in healthcare service delivery. The Bayesian Classifier Techniques has been used to diagnosis, predict diseases in healthcare but not in the area of the combination of the identified common five diseases (Malaria, Typhoid, Yellow Fever, Cholera and Tuberculosis) considered in this paper. Arbaiy et al (2017) used expert system technique to develop a Web-based Community Center for Healthcare Diagnosis system. Web based technology is used as a platform to disseminate the information to users in order for them to optimize the information appropriately. The system helps the public to identify the type of diseases based on the symptoms given and early detection can allow prevention of serious illness. The system also provides information related to arthritis, thalassemia and pneumococcal and this was implemented using Adobe Dreamweaver CS6, Xampp, and Adobe Photoshop CS6. De Silva and Jayamanne (2016) employed a Bayesian framework to construct a Web-based decision support system for medical diagnosis.

The proposed system identifies diseases by analyzing symptoms or by analyzing medical test results. The system detects different types of diseases that people suffer in their day-to-day lives (general diseases) with an average detection accuracy of 92.56%. System also detects complex diseases (e.g.: heart disease - 83.67%, breast cancer - 80.98%, liver disorders - 79.43%, lung cancer - 71.00%, primary tumor - 78.02%, etc.) based on the analysis of the medical test results. Chang, Fan, Lo, Hung, & Yuan (2015) presented an inference model based on an ontology and a Bayesian network to infer the possibility of becoming depressed. The prototype was implemented using a mobile agent platform as a proof-of-concept in the mobile cloud. This framework exploits the ontology technique to construct a DOM and it utilizes a BN to build a model to infer the probability of becoming depressed. Adetunmbi, Oguntimilehin & Falaki (2012) developed a web based system for malaria diagnosis and therapy using rough set as a machine learning technique to label sets of malaria fever symptoms collected to generate explainable rules for each level of severity and appropriate therapy provided. The labelled database was divided into five cases of malaria and the classification accuracy on training dataset is 100% while that of testing data set is 94%. This system was developed using HTML and PHP as front end and MYSQL as backend.

3. METHODOLOGY AND FRAMEWORK

The researchers visited Yaba College of Technology (YCT) medical center and some other clinics in Yaba environs in order to obtain comprehensive information about diseases and their symptoms and how the medical diagnosis are been carried out. It was also discovered from investigation that:

i. The medical records of patients are kept manually, these are tedious to retrieve and not safe nor secured.
ii. Diagnosis are based on individual doctor's skills and personal experiences which can be characterized to be inaccurate and imprecise.
iii. Most often patient wait on long que for several minutes even hours before they are attended to.
iv. Method of diagnosis is the traditional method where the physician asks the patient for symptoms being experienced, examines the patient physically, checks the medical history or case note and laboratory test result manually
v. Also from the review of related works, it was discovered that the existing medical diagnosis system focused on one disease.

This paper aimed to develop a web based medical diagnosis system using Bayesian Classifier Technique. HTML (Hypertext Markup Language) was used to develop the front end and PHP (Php Hypertext Preprocessor) for the backend of the system. Structured Query Language (SQL) was used for the querying of the database. The Visual Analogue Scale (VAS) was used to present the result of the Bayesian classifier into their level of severity (no disease, mild disease, Moderate disease and severe disease). Twenty six datasets features (symptoms) were considered based on our findings which focused on five common diseases (malaria, typhoid fever, cholera, tuberculosis and yellow fever) among the students of YCT. The system accepts twenty six datasets features from the user which were used in determining the likelihood of any of the five common diseases among the students. The system is able to predict whether a disease is positive or not for a new set of data by using different measures obtained from conducting various test per disease. The system classified the symptoms using Bayesian Classifier as stated below.

Let the probability of event P describes set of symptoms (S) i.e. S1, S2, S3....Sn. The symptoms are assumed to be independent, then, P (S1, S2, S3 ...Sn) = P(S1)P(S2)P(S3)...P(Sn). The posterior probability of the disease being positive could be computed. The probability of a particular disease is calculated using:

$$P(+|S1,S2,S3...\ Sn) = \frac{P(S1,S2,S3\ldots Sn|+) * P(+)}{P(S1,S2,S3\ldots Sn|+)* P(+) + P(S1,S2,S3\ldots Sn|-) * P(-)}$$

Where P(S1, S2, S3, S4....Sn|+) is the probability of confirming if the presence of a symptom can cause a disease while P (+) is the probability that the disease is present and P(–) is the probability that the disease is not present (De Silva and Jayamanne, 2016).

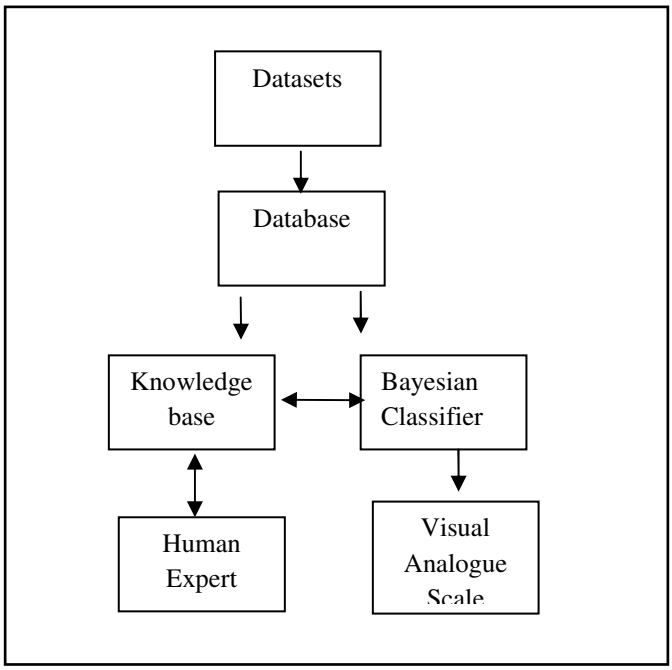

Figure 1: Block Diagram of System Workflow

4. DATA PRESENTATION

Table 1 showed the list of Twenty six datasets that is the symptoms available on the Diagnosis interface where the user select "Yes" to denote that the symptom exist or "No" to denote that the symptom did not exist. Also Table 2 showed the list of five diseases that were considered with their various symptoms, this is used to classifier the symptoms to the diseases their matched.

Table 1: Datasets Features

S/NO	SYMPTOMS	S/NO	SYMPTOMS
1	High_fever	14	Headache
2	Nausea	15	Vomiting
3	Abdominal_pain	16	Diarrhea
4	Muscle_pain	17	Dehydration
5	Fever	18	Watery_diarrhea
6	Convulsion	19	Constipation
7	Fatigue	20	Delirium
8	Chills	21	Loss_of_appetite
9	External Bleeding	22	Yellow_skin_eyes
10	Internal_bleeding	23	Liver_injury
11	Deceased_urination	24	Liver_failure
12	Chronic_pain	25	Cough_blood
13	Chest_pain	26	Weight_loss

Table 2: Five Diseases Based on the Symptoms

MALARIA	TYPHOID FEVER	CHOLERA	YELLOW FEVER	TUBERCULOSIS
Fever	Abdominal pain	Diarrhea	Fever	Fever
Chills	Headache	Vomiting	Chills	Chest_pain
Headache	High_fever	Dehydration	Fatigue	Fatigue
Nausea	Internal_bleeding	Fatigue	Chronic_pain	Cough_blood
Vomiting	Diarrhea	Delirium	Nausea	Weight_loss
Delirium	Constipation	Abdominal_pain	External_bleeding	Loss _of_appetite
Fatigue	Loss_of_appetite	Fever	Vomiting	
	Fatigue	Weight_loss	Liver_failure	
	Diarrhea	Deceased_urination	Yellow_skin_eyes	
	Delirium	Watery_diarrhea	Convulsion	
			Liver_injury	
			Delirium	
			Abdominal_pain	
			Muscle_pain	
			Headache	

4.1 Scoring and Interpretation

Visual Analogue Scale was used for both the input and output to generate a random number for the numeric data. The scale used for the input is from zero to one, the zeros indicate that the symptoms did not exist while ones indicate that it exist see Figure 2 below. The result from the Bayesian classifier is based on the distribution of diseases and VAS described the diseases as no disease, mild, disease, moderate disease, or severe disease. The following cut points as it in VAS were used: no disease (0–4 mm), mild disease (5-44 mm), moderate disease (45–74 mm), and severe disease (75–100 mm) (Gould, 2001) see Figure 3 below.

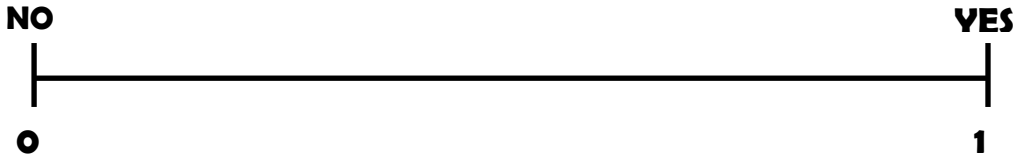

Figure 2: Input Visual Analogue Scale

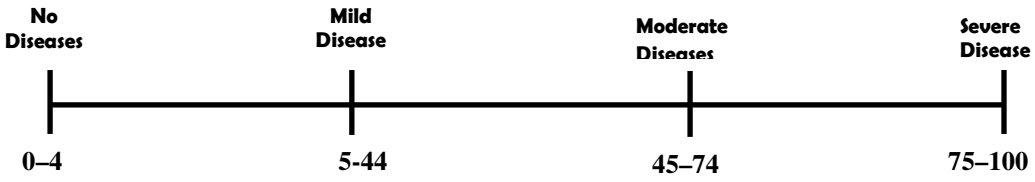

Figure 3: Target Output Visual Analogue Scale

Table 3: Target output percentage

Diseases Classifier	Percentage Classification	Class Number
No diseases	0-4.99%	1
Mild diseases	5-44.99%	2
Moderate diseases	45-74.99%	3
Severe diseases	75-100%	4

5. IMPLEMENTATION AND TESTING

The system was developed using HTML (Hypertext Markup Language) for the front end and PHP (Php Hypertext Preprocessor) for the backend. Structured Query Language (SQL) was used for the querying of the database. The system was implemented based on the data in Table 1 and Table 2. The figures below showed the splash screens after launching of the program. Figure 4, Figure 5 and Figure 6 showed the splash screen of Registration Interface, splash screen of Diagnosis Interface and splash screen of Diagnosis Report respectively. In the registration interface, the user register by supplying his personal data before he has access to diagnosis interface where symptoms are selected by clicking yes or no.

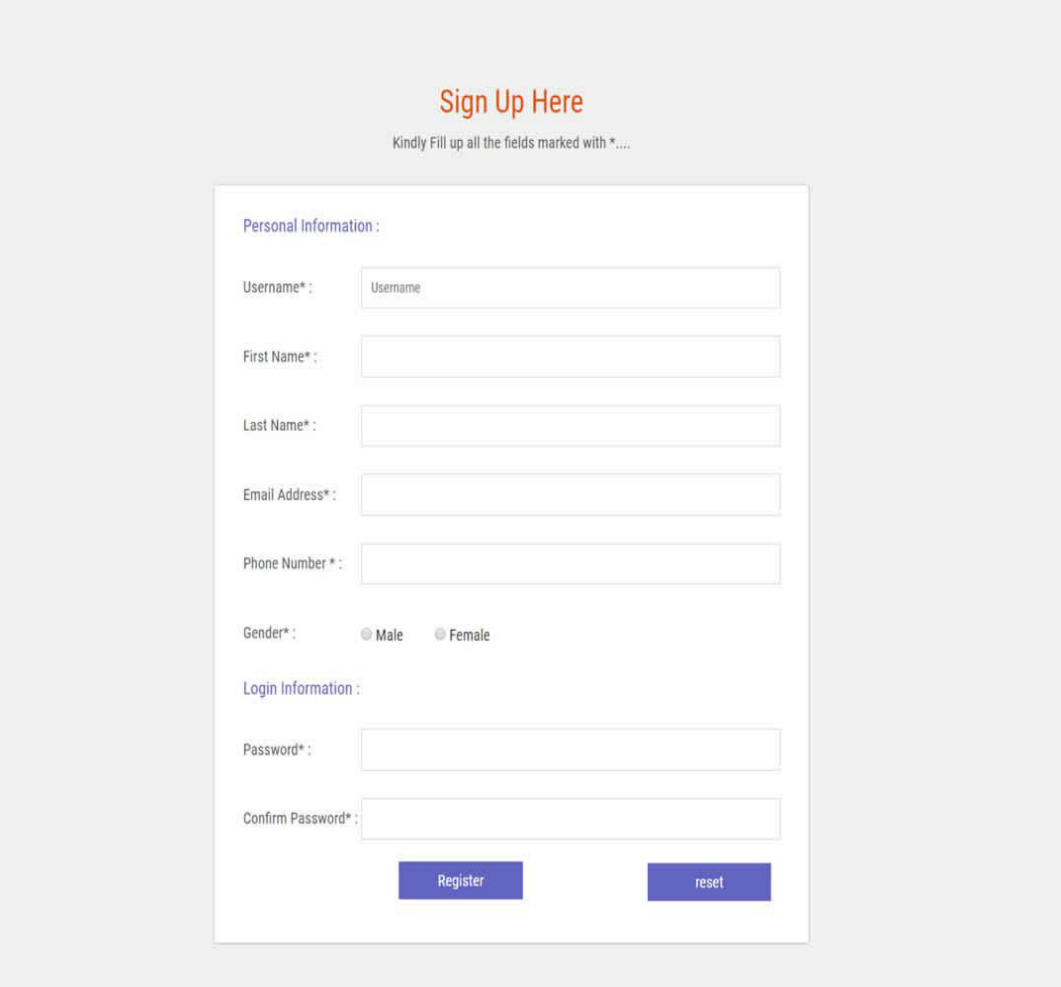

Figure 4: Registration Interface

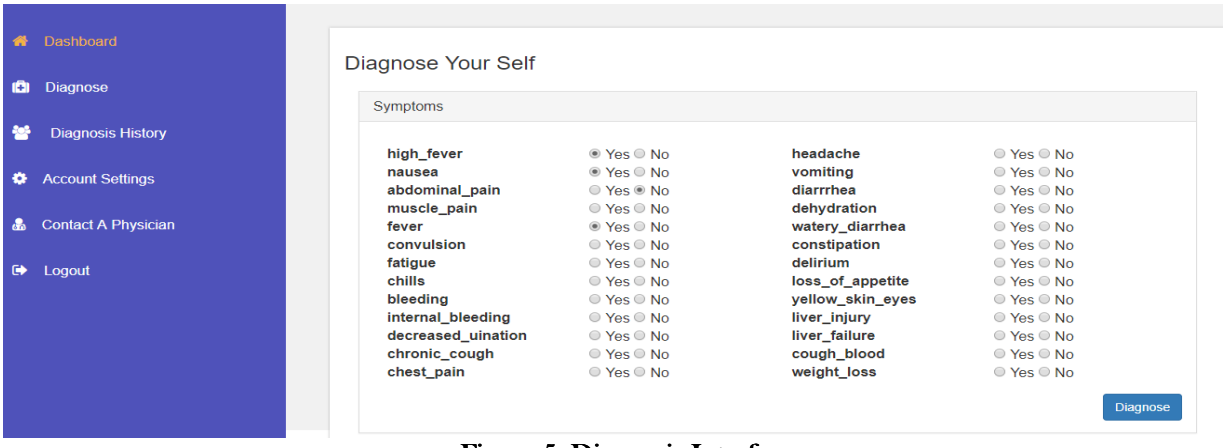

Figure 5: Diagnosis Interface

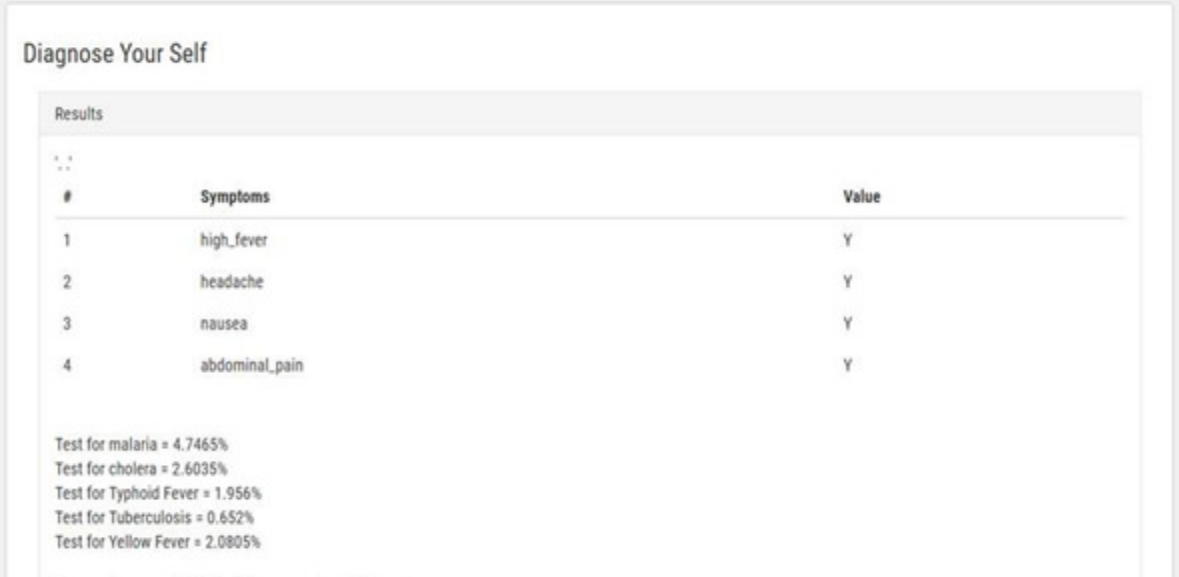

Figure 6: Diagnosis Report

6. DISCUSSION OF RESULTS

Fifty data were generated using the datasets in table 1. The severity of the diseases were classified into No disease, mild disease, moderate disease or severe disease based on the visual analogue scale for the five common diseases considered in this paper. The output having the highest class number in the diseases considered will be recommended for treatment based on the dataset features for diagnosis.
Table 4 and Table 5 depict the result for ten users based on the symptoms selected from the diagnosis interface. Yes option is denoted by 1 and No option is denoted by 0 on the table. Table 6 depict the diseases recommended for the treatment and the severity if the diseases for the ten users.

Table 4: Sample Datasets A

User	High Fever	Nausea	Abdominal pain	Muscle pain	Fever	convulsion	fatigue	chills	External bleeding	Internal bleeding	Decreased urination	chronic cough	Chest pain
User 1	0	1	1	1	1	1	0	0	1	1	0	0	1
User 2	1	1	1	1	1	0	1	0	1	1	1	0	1
User 3	0	1	1	0	1	1	1	1	1	0	0	1	0
User 4	0	0	0	1	0	1	1	0	0	0	1	0	1
User 5	0	1	0	1	1	1	0	1	0	1	0	0	1
User 6	0	0	0	0	1	1	0	1	0	0	1	1	0
User 7	1	1	0	1	0	0	1	0	0	1	1	1	0
User 8	0	0	1	0	0	1	1	1	1	0	1	1	0
User 9	1	1	1	1	1	0	0	0	1	1	1	1	0
User 10	0	1	0	0	0	1	0	0	1	1	1	1	1

Table 5: Sample Datasets B

User	headache	vomiting	diarrhea	dehydration	Watery diarrhea	constipation	delirium	Loss of appetite	Yellow Skin eyes	Liver injury	Liver failure	Cough blood	Weight loss
User 1	0	1	0	0	1	1	0	1	0	1	0	1	0
User 2	0	1	1	1	1	1	1	0	0	1	1	1	0
User 3	0	0	0	0	0	1	0	0	1	1	1	0	1
User 4	0	0	0	1	1	1	0	1	1	1	0	1	1
User 5	0	0	0	1	1	1	0	0	1	1	1	1	1
User 6	1	1	0	0	0	0	0	1	1	1	0	1	0
User 7	0	0	1	0	1	0	0	1	1	0	0	0	1
User 8	1	1	0	0	1	0	1	0	1	1	0	0	0
User 9	0	1	1	1	1	0	0	1	0	0	1	0	1
User 10	1	1	0	1	0	0	0	0	0	1	0	1	0

Table 6: Sample Result C

User	malaria	Typhoid	Yellow Fever	Cholera	Tuberculosis	Recommended Treatment	Class number	Level of severity
User 1	1.1666667	1.5555556	2.3333333	1.1666667	1.5555556	2.333333333	2	mild
User 2	2.6388889	3.6944444	3.6944444	3.6944444	2.1111111	3.694444444	4	severe
Use 3	1.4444444	1.0833333	3.6111111	1.0833333	1.0833333	3.611111111	4	severe
User 4	0.3611111	1.0833333	2.1666667	1.8055556	1.8055556	2.166666667	2	mild
User 5	1.25	0.8333333	3.3333333	1.25	1.6666667	3.333333333	3	moderate
User 6	1.2222222	0.6111111	2.1388889	0.6111111	0.9166667	2.138888889	2	mild
User 7	0.6666667	1.6666667	2	1.6666667	1	2	2	mild
User 8	1.8055556	1.4444444	3.25	1.8055556	0.3611111	3.25	3	moderate
User 9	1.3333333	2.2222222	3.1111111	3.1111111	1.3333333	3.111111111	3	moderate
User 10	1	0.6666667	2.3333333	1	0.6666667	2.333333333	2	mild

7. CONCLUSION

This paper reviewed previous works on medical diagnosis system and developed a Web-Based Medical Diagnosis System using Bayesian Classifier Techniques. The system diagnosed patient on real-time as the doctor does to the patient with or without the presence of the physician at any time. The System automates the whole process of patient's diagnosis with accuracy and present the severity of the diseases. The VAS was used to evaluate the result of the Bayesian classifier into their level of severity (no disease, mild disease, moderate and severe disease) and the result is used to determine further recommendation for the patient. This system will assist the health workers and physicians in discharging their duties at appropriate time to safe life. It was observed that the research objective were achieved and the program will be deployed for usage within and outside the college.

REFERENCES

1. Adetunmbi, A. O., Oguntimilehin, A., & Falaki, S. O. (2012). Web-Based Medical Assistant System for Malaria Diagnosis and Therapy. Journal of *Computer Science and Telecommunications*, *33*(1).
2. Arbaiy, N., Sulaiman, S. E., Hassan, N. and Afip, Z. A. (2017). Integrated Knowledge Based Expert for Disease Diagnosis System. International Research and Innovation Summit (IRIS 2017). Available online at http://www.iopscience.iop.org/article/10.1088/1757-899X/226/1/01097/pdf. Retrieved May 25, 2018.
3. Chang, Y. S., Fan, C. T., Lo, W. T., Hung, W. C., & Yuan, S. M. (2015). Mobile Cloud-based depression diagnosis using an ontology and a Bayesian network. *Future Generation Computer Systems*, *43*, 87-98.
4. De Silva, N. T. and Jayamanne, D. J. (2016). Computer Aided Medical Diagnosis using Bayesian Classifier – Decision Support System for Medical Diagnosis. International Journal of Multidisciplinary studies (IJMS), 3(2). 91-97.
5. Fatumo, S. A., Adetiba, E., & Onaolapo, J. O. (2013). Implementation of XPERTMALTYPH: An Expert System for Medical Diagnosis of the Complications of Malaria and Typhoid. *IOSR Journal of Computer Engineering (IOSR-JCE)*, *8*(5), 34-40.
6. Gould D. (2001). Visual Analogue Scale (VAS). Journal of Clinical Nursing. 10(1). 697-706. Available Online at http://www.physio_pedia.com/visual_analogue_scale. Retrieved May 25, 2018.
7. Indhumathi, S., & Vijaybaskar, G. (2015). Web Based Health Care Detection Using Naïve Bayes Algorithm. International *Journal of Advance Research in Computer Engineering and Technology (IJARCET)*, *4*(9), 3532-3536.
8. Raj, T. M., & Prasanna, S. (2013). Implementation of ML Using Naïve Bayes Algorithm for Identifying Disease-Treatment Relation in Bio-Science Text. *Journal of Applied Science, Engineering and Technology*, *5*, 421-426.
9. Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford University Press, USA.
10. Suchánek, P., Marecki, F., & Bucki, R. (2014). Self-Learning Bayesian Networks in Diagnosis. *Procedia Computer Science*, *35*, 1426-1435.