Minimizing Symptom-based Diagnostic Errors Using Weighted Input Variables and Fuzzy Logic Rules in Clinical Decision Support Systems

Anietie Ekong[1], Henry Odikwa[2], Otuekong Ekong[3]
[1]Department of Computer Science, Akwa Ibom State University, Nigeria, anietieekong@aksu.edu.ng
[2]Department of Computer Science, Abia State University, Nigeria, odikwa@yahoo.com
[3]Department of Computer Science, Akwa Ibom State University, Nigeria, otuekong@gmail.com

ABSTRACT

This study covers an approach towards minimizing Disease diagnostic errors using weighted input variables and Fuzzy Logic rules with multiphase diagnostic engine. The weights were applied because different symptoms may have different degrees of importance in different diseases. This is to ensure that recommendations for disease confirmation based on symptoms return good percentage of true positive and true negatives. The study creates an enhanced, accurate and precise system for medical diagnosis even when only the symptoms are considered. In order to evaluate the model, four categories of diagnoses were carried out without using the model at the first instance and using the model at the second instance with 50 patients done at 4 different diagnostic instances. The true positive (TP) and the False negative statistics were obtained from where the false positive rate (TPR) or sensitivity and false positive rate (FPR) were derived. The graph of TPR vs FPR was plotted from where the quality of diagnoses could be gotten from the Receiver Operating Characteristics (ROC) space. The result shows that sensitivity, which is the ability of a test to correctly identify those with the disease or True Positive Rate, and specificity, which is the ability of the test to correctly identify those without the disease also called True Negative Rate TNR stood at 87% and 86% respectively using the developed model and the same parameter yielded 72% and 56% respectively without using the model. The result also shows that the false positive rate (FPR) which indicates the degree of false alarm is 19% using the new model while it is 44% without using the model. This result shows that the likelihood of making wrong clinical diagnostic decisions is much lower with this approach.

Key words: Clinical Decision, Disease Diagnostic Errors, Diagnostic Algorithms, Fuzzy Logic, Weighted Input.

1. INTRODUCTION

One of the major problems with respect to human's health is proper diagnoses. With the shortage of medical experts worldwide, the risk of being improperly diagnosed is on the increase. To maintain the few experts, hospitals are charging exorbitant fees from patients and this further complicates the health conditions of people especially in low income countries. It is necessary to find a better solution by ensuring that diagnoses are not based on the traditional approaches that have been in existence for ages which are tainted by imprecise, delayed or outrightly wrong diagnoses. There are several existing approaches to improve on the sensitivity, specificity and accuracy of disease diagnoses but they are still well away from perfection and might have only added a very small improvement to the human-only approach to diagnoses hence the need to continuously improve on the approaches. According to Graber, Franklin and Gordon [3], a diagnostic error is a consequence of a missed, inappropriate, delayed or wrong diagnosis. Medical diagnosis is the classification of symptom presentations to make informed decisions about the disease that might have given rise to such symptoms for the purpose of treating or proffering solutions to such underlying causes. Due to the fact that many signs and symptoms are non-specific, medical diagnoses could be a daunting task as many visible symptoms may have resulted from different disease' types. This can lead to the mapping of symptoms presented to a wrong disease. In order to minimize the case of wrong diagnoses, experts have always used the idea of grouping signs or symptoms that are indicative of a particular disease or condition. Due to the low ability of many people to repeat same process precisely, different diagnoses by different experts at different times occurs frequently.

A major drawback to the traditional method of medical diagnoses is the inaccuracy of such diagnoses which has led to many loss of lives among other undesirable consequences. With the improvement in computers and with the advent of Artificial Intelligence, computers have come handy to improving the ways that medical diagnoses are handled and this has led to increase in accuracy, reduced diagnostic time and considerable reduction in the number of deaths. Fuzzy logic and its rules have been of profound importance in clinical support systems, and with the use of Fuzzy inference which is the mapping from a given input to an output using fuzzy logic, this has led to more precise and accurate diagnoses hence helping in administering the proper treatment as a wrong diagnosis is worse than no diagnoses. This problem is aggravated by lack of adequate medical personnel especially in low income countries. The importance of this approach is to help medical practitioners

in making good clinical decisions.

According to Preety, Aman and Deepti [9], there are a number of properties in fuzzy set theory with a number of facilities that make it suitable for clinical decision support systems. Though, there are many systems for diagnoses, most of them focus on single diseases prognosis. Weighted inputs and fuzzy logic rules is chosen as the artificial intelligence tool employed in the proposed model because fuzzy logic is an efficient computational method in handling degree of uncertainties and the weighted input ensures that different input variables(weighted symptoms) have different importance towards the diseases. The model is such that, given the symptom set with their degree of importance and assuming that other associated parameters are similar, the probability of it being mapped to a particular disease repeatedly is one(1). In most cases, the symptoms are only-suggestive of a disease or diseases, so the degree to which the symptoms sets presentations point to a particular disease must be taken into account and must be the major determining factor in disease diagnoses.

The weighted input and rule based Fuzzy Clinical Support System will accept a set of symptoms and maps it to a particular disease based on their weights, inference engine and the knowledge base. Fuzzy logic has type-I and type-II. In type 1 fuzzy logic, the exact degree of achieving membership of an object must be determined. Though the type-2 fuzzy logic was later introduced which was described as an extension general fuzzy set with a new dimension called the footprint of uncertainty, type-2 fuzzy set is characterized by a fuzzy membership function while tye-1 fuzzy membership function is crisp and are excellent in scenarios of certain membership functions. In this research, the rules based for type I fuzzy logic is used because the degree of membership for input variables (symptoms) is known. Since it is mostly not possible to have a 100% certainty of a particular disease based on the symptoms presented by people, partly as a result of varying body chemistry among other factors, experts are unable, in a general sense, to determine the exact degree to achieve the characteristics of its membership sets and there must be a way of taking care of the inexactness.

The aim of this research is to remarkably reduce diagnostic errors in clinical decision support.

2. MOTIVATION

The burden of making accurate clinical decision is a major concern in healthcare delivery as so many factors play varying roles, some of which mitigate against preciseness of such decisions. These factors include but not limited to; the similarity of symptoms of many diseases, the variation in the biological response by different people, the nature of their living and working environment and the background of the patients. These have, to some extent, made clinical decisions to be subjective as they are being affected by the opinion of different experts. Having large number of variables to analyze before diagnoses is a daunting task to health care professionals and this has led to wrong diagnoses or delayed diagnoses with heavy price to pay by the patients and the society. To this end, an appropriate tool to assist in analyzing the variables will surely be of immense importance and will lead to a more accurate diagnoses with its attendant benefits and such analysis can be repeated for as many times as possible since the computers have the abilities for iterative task with extreme precision. This ensures that similar symptom sets could always be mapped to the same result using the same knowledge base and inference engine. Motivated by the above, a model using a weighted input and fuzzy logic rules to support clinical decision has been developed and quality of diagnoses analyzed with the aid of Receiver Operating Curve (ROC).

3. RELATED LITERATURE

"Accessible and safe healthcare is essential to achieving universal health coverage" World Health Organization [11]. For instance in HIV diagnoses, the World Health organization (WHO) posited that the diagnostic errors were mostly from humans ranging from suboptimal testing strategies; testers having difficulty interpreting weak reactive results, poor management to laxity in the supervision of testers and quality of systems in the laboratories, World Health Organization [13]. Although the levels of misdiagnosis reported are low, by estimates, the large volume of tests conducted each year - about 2.3 million new infections globally in 2018, World Health Organization [12], the implication of a wrong diagnoses could result in the misdiagnosis of a large number of those people being classified as HIV positive with its grave consequences.

Fuzzy Expert System has proved its usefulness in medical diagnosis for the quantitative analysis and qualitative evaluation of medical data, consequently achieving the correctness of results, Manish and Sedamkar [7]. A crisp set is a collection of objects whose values are binary in nature, which could answer whether the object belongs to a particular set or not. This is called crisp or Boolean logic. In fuzzy logic, an object belongs to a particular set to a certain degree and the degree to which it belongs to this set is termed its membership value. This membership value varies from 0 to 1 and the function that defines how each input is mapped to membership value (or degree of membership) between 0 and 1 is membership function Preety et al. [9]. Considering a classical set expressed as $S = \{x \mid x > n\}$ where n is some integer. A fuzzy set is an extension of a classical set. If U is the universal set in consideration comprising of elements denoted by y, then a fuzzy set S in U is defined as a set of ordered pairs: $S = \{y, m_S(y) \mid y \, \varepsilon U\}$ $m_S(y)$ is called the membership function (or MF) of y in S and MF maps each element of U to a membership value between 0 and 1.

Rule based Fuzzy logic expert systems are of immense importance to medical diagnoses and ensures a more precise evaluation of symptoms and other data, the method is straightforward and fast and could be used by anybody and is particularly useful where there are no sufficiently qualified medical practitioners and in clinical decision support systems.

Johnston, Brian and Haynes [5] reviewed the evidence from controlled trials of the effects of computer-based clinical decision support systems (CDSSs) on clinician performance and patient outcomes, the result showed improvements in clinician performance using a CDSS. They concluded that there was strong evidence which suggested that some CDSSs can improve physician's performance. According to Ahmed, Sherif and Ahmed [1], Artificial Intelligence have been used in medical applications and other domains and, many decision support systems (DSSs) like Aaphelp, Internist I, Mycin, Emycin and others have been designed and they assist in diagnosis and treatment of different diseases".

Das, Guha, and Dutta [2] combined fuzzy rule-based and intuitionistic fuzzy based inference systems for better and more realistic representation of uncertainty of the medical diagnosis problem and for more accurate diagnostic result. They used the modeling of antecedent part of the rules consisting of linguistic assessments of the patients' symptoms provided by the doctors/medical experts with their corresponding confidence levels, by using generalized fuzzy numbers and the modeling of consequent part, which reveals the degree of association and the degree of non-association of diseases, intuitionistic fuzzy aggregation operator in inference process and the application of relative closeness function to find the final crisp output for a given diagnosis to get a valid conclusion for a set of patients.

A classifier that tackles the problem of determining the risk of a patient suffering from a cardiovascular disease within a certain time frame is possible and has been developed according to José, Mikel, Aranzazu, Antonio and Humberto [6]. They provided both a diagnosis and an interpretable model explaining the decision where doctors are able to analyse the usefulness of the information given by the system. They were able to show the suitability of the system to deal with this medical diagnosis classification problem and that the performance of the new method was statistically better than the ones obtained with the methods considered in the comparison. They were able to enhance both the total number of correctly diagnosed patients, and the classifier had the ability to correctly differentiate patients of the different risk categories. They concluded that the system was a suitable tool to face the medical diagnosis of cardiovascular diseases since it obtained a good classification rate and it also provided an easily understood and interpretable model by doctors.

Hossein, Omid, Javad, Gholamali and Mohammad [4] studied the use of thermal image scanning to help, where it is difficult to determine with certainty, that the observed pattern in the image of the breast is benign or malignant and arrived at the conclusion that clinical support tool can assists physicians in distinct diagnostic tasks by providing a relatively accurate decision support tool in arriving at an accurate diagnosis based on the type of imaging pattern as a response from the computer model hence eliminating or at least reducing the possibility diagnostic error in misrepresenting the observed condition of benign as malignant or vice versa.

The reformulated of diagnostic procedure in terms of fuzzy relation equations, from which results are retrieved and extended was done by Sanchez [10]. He formed a truth-possibility index which serves pattern-matching purposes, in addition to the usual possibility and necessity measures and created a biomedical application, in which medical knowledge is expressed in a rule form and together with fuzzy propositions in the antecedent and the s um total of all these measures helped in providing the necessary medical diagnosis assistance and reduce diagnostic error.

Muhammad, Saleem, Yi and Muhammad [8] studied an advanced approach to linguistic intuitionistic fuzzy variable through application of cubic set theory, established a series of weighted aggregation operators under linguistic intuitionistic cubic fuzzy information and their fundamental properties and showed their relationship and having developed a multi-criteria decision-making algorithm, they demonstrated that the algorithm performs better than pre-existing aggregation operators.

Yang, Li, Garg and Qi [14] developed a decision support algorithm based on spherical normal fuzzy (SpNoF) set for selecting antivirus masks. They posited that there is great difficulty for decision makers to properly select the right masks and by analyzing the accuracy function, improving on operational rules and their properties, working on the SpNoF Bonferroni mean operator and the weighted Bonferroni mean operator, they were able to established a multi-criteria decision-making method that proved the efficacy of decision support systems.

According to Giovanni, Valeria, Donatella, Stefania and Maria[15], clinical decision support systems (CDSS) play an increasing crucial role in improving the quality of the therapeutic and diagnostic efficiency in health care and fuzzy logic provides an effective means for dealing with uncertainties in the health decision-making process.

According to Manish and Sedamkar [7], in medical diagnosis, decision making are affected by many variables which results in differing opinions by the experts. They posited that due to many uncertain factors, disease diagnoses is hard and hence a more accurate tool would be of great help to assess all the risk factors so as to get more accurate results.

4. RESEARCH METHODOLOGY

The approach adopted in this research is weighted input and fuzzy logic rules. It is a multiphase approach which helps to minimize or eliminate errors with the aid of a clinical decision support system. Phase I is based on symptoms and the output is a probable disease and the degree to which it is probable makes the disease an input candidate for phase II. Phase II is the laboratory test or any further test other than

symptom based assessment to confirm the probable disease in phase I. In phase I, each of the inputs which are the symptoms of the diseases are weighted based on their degree of importance for the proper classification of the disease. The data is collected by different experts who assigned different weights to the inputs and at the end, the average weight for any symptom was derived from the formula

$$WS = \left(\sum_{i=1}^{n} Wi\right)/N \quad (1)$$

Where WS is the average weight of the symptom; W is the weight of each of the symptoms and N is the total number weighted symptoms.

The subjects are experts in the field of medicine and they are not expected to, as a matter of necessity, be knowledgeable about computer systems or fuzzy rules set. This study considered a set of diseases E and their respective symptoms as obtained from the experts E = {D$_1$, D$_2$,...., D$_n$}; D represent the disease, the symptoms S of the diseases in E are also defined as S(D$_i$)= {S$_1$, S$_2$,.... S$_n$};

The weighted inputs (symptoms) form the knowledge base and the inference engine is built according to the algorithm in Section IV-*C*. The presented symptoms from patient are compared with the weighted symptoms in the knowledge base to appropriately map such symptoms to the disease. This ensures accurate diagnoses and help reduce the work load for medical practitioners with advantages to all the parties involved. The output of phase I is based on the rules that combined to produce an output which is the probable disease. The disease that is the mostly likely is suggested as the probable disease and this is determined by the return of the disease with the highest number of symptom sets that the patients presented in relation to the diseases with their weighted symptoms in the knowledge base. That is, if two diseases return as the probable disease, the one with the highest symptom weight is considered more likely, then the

system progresses to refer to Phase II for out the system laboratory test for confirmation or otherwise of the suspected diseases. The result in phase I is compared with the result in phase II and the number of True positive(TP), True Negative (TN) and False Positive(FP) are obtained and from there, the True Positive Rate(FPR), False Positive Rate(FPR) and True Negative Rates (TNR) are computed and the Receivers operating characteristic curve is used to determine the quality of the system. Figure 4 shows the diagnostic phases.

In consideration of the above, if a prediction outcome is D and the actual or confirmed value as D, then it is called True Positive; but if the confirmed or actual value is, say M then it is said to be a False Positive i.e raising of false alarm. A true negative occurs when both the prediction outcome and the actual value are M while a false negative occurs when prediction outcome is M while the actual value is D. Many data from patients were used to test the system and to see if the diagnosed would have been different using different set of medical personnel other than the ones we used as experts. In this research, a true positive is when the diagnosed disease in phase I is the same as the confirmed disease in phase II, a false positive is when the diagnosed disease diagnosed in Phase I is different from the confirmed disease after phase II, that is the patient does not actually have the disease as suggested earlier in Phase I, a true negative is when the outcome in phase I doesn't suggest the disease and the confirmatory test in phase II is also negative while a false negative, on the other hand, is when the diagnosed disease in Phase I doesn't suggest the disease whereas the value in phase II points to the disease, assuming that phase II could be trusted in absolute terms. Figure 1 illustrates the conditions for each outcome while Figure 2: Illustrates the weighted input (symptoms) or Membership Functions & Confirmatory test Outcome is
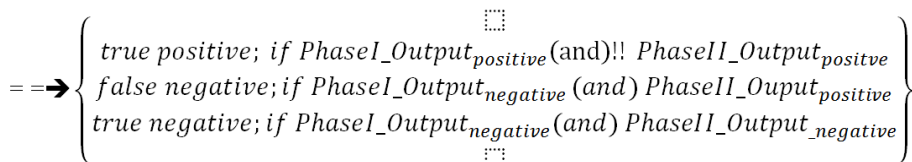
$$= \Rightarrow \begin{cases} true\ positive;\ if\ PhaseI\_Output_{positive}\ (and)!!\ PhaseII\_Output_{positve} \\ false\ negative;\ if\ PhaseI\_Output_{negative}\ (and)\ PhaseII\_Ouput_{positive} \\ true\ negative;\ if\ PhaseI\_Output_{negative}\ (and)\ PhaseII\_Output_{\_negative} \end{cases}$$
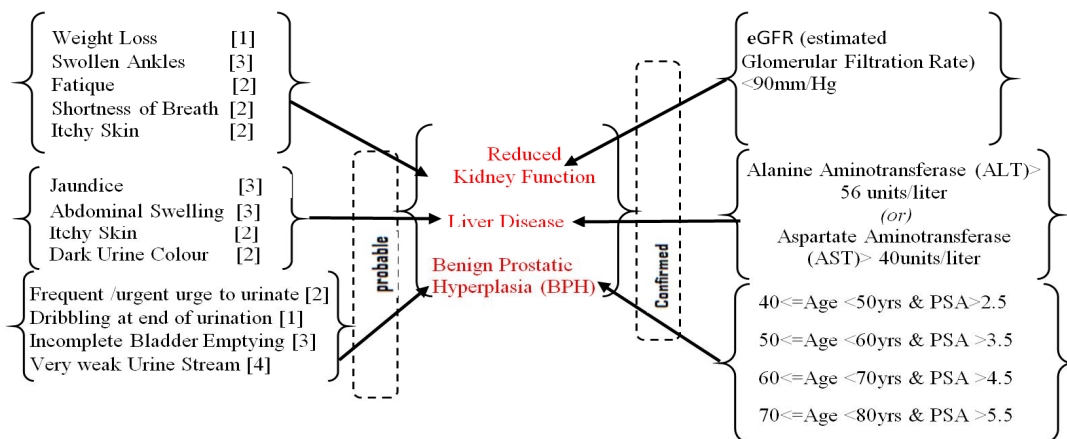
Figure 1. Conditions for the prediction outcome.

Figure 2. Illustration of the Weighted input (symptoms) or Membership Functions &Confirmatory test.

Where eGFR=186 x (Creatinine/88.4)$^{-1.154}$ x (Age)$^{-0.203}$ x (0.742 if female) x (1.210 if black)
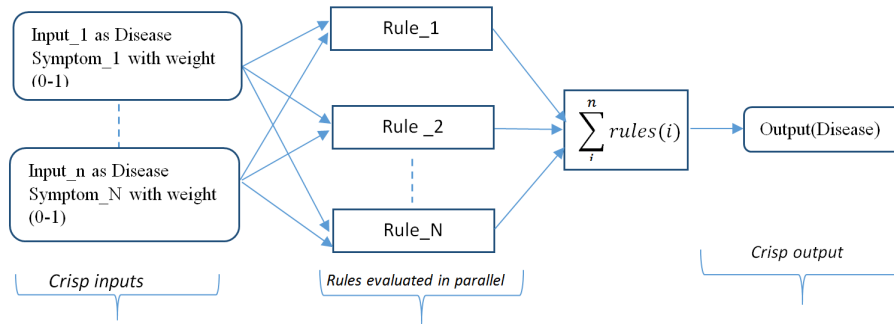

Figure 3. *Illustration of the Fuzzy Logic Rules.*

A. FUZZY RULES

The next step is the development of fuzzy rules. Here the rules were developed with the assistance of medical experts consisting of three (3) medical personnel. The study chose only the relevant rules among many others in the knowledge base and a rule triggers if any of the precedent parameters, patient' symptoms set, intersects with the disease symptom Set and evaluate to true (1)and false (0) otherwise; the output is the summation of the rules as illustrated in Figure 3. Put it differently, a rule is activated if any of the patient's symptom is associated with any of the disease symptoms. Here, the disease with the highest number of patients symptoms in the disease symptom set is the probable disease. Where more than two diseases have the same symptom sets, the weight is put into consideration and the one with the highest weight is taken to be the probable disease. The fuzzy inference system takes in the input as symptoms and produce crisp output which is the name of the probable disease in phase I and based on that, the output is fed into phase II where the appropriate laboratory test is carried out to dispel or confirm the probable disease in Phase-I. Figure 3 is an illustration of the fuzzy logic rules.

B. DEFUZZICATION & INFERENCE ENGINE

The defuzzifier produces a crisp output from the inference engine whose input was a fuzzy set since crisp values are need in most real world scenario. The inference engine is a component of the system that draws conclusion or derived further or new information from existing data after the application of the necessary logical rules to the knowledge base, it is modelled after domain's expert reasoning. In this case, it is made to draw appropriate conclusions which is the probable diseases based on the symptoms associated to the patients and the ones associated with the disease. Fuzzy inference is the mapping from a given input to an output using the theory of fuzzy sets and in this case the mapping of the presented symptoms to the disease. We use the Mamdani Inference type in this research. This inference engine uses a forward chaining approach to search the knowledge for the symptoms of the disease. Forward chaining generally starts with the known facts and asserts new facts and is one of the techniques in supervised learning. The weight is used in creating the strength or importance of different input variables and rules that have a common conclusion.

C. ALGORITHMS TO BUILD THE KNOWLEDGE BASE

Step 1: input diseases into the knowledge base
Step 2: loop through all disease in the knowledgebase and
> {
> *Step 3:create an empty symptom list for each disease in [1]*

Step 4: input symptoms to all diseases in [3]
> *Step 5: create an empty symptom weight list for each symptom in [4]*

Step 6: input weight for each symptom in [4] as it pertains to all the diseases in [1]
Step 7:return all disease with its symptoms list and associated weight
> }

D. PSEUDOCODE TO BUILD A KNOWLEDGE BASE

SUB Build_Disease_KnowledgeBase()
  FOR i=1 to diseasecount
    input diseases in the knowledge base
    SymptomLlist(i)=[ ]
  NEXT I
  FOR i=1 to DiseaseCount
    FOR j=1 to SymptomCount
    *SymptomList[i]=SymptomList[i]&Symptom_*
    &SymptomWeight
    NEXT j
  NEXT i
SAVE(All_DiseaseList_With_WeightedSymtoms)
END SUB

E. ALGORITHMS TO ACCEPT PRESENTED SYMPTOMS BY PATIENTS

Step 1: initialize presented symptoms list to nil
Step 2: loop through all symptoms as presented or observed
> {
> *Step 3: add the observed symptom to observed symptom list*
> *Step 7:return all observed symptoms list*
> }

F. PSEUDOCODE TO ACCEPT PRESENTED SYMPTOMS BY PATIENTS

Sub Input_Patients_Symptoms()
  Patients_Symptoms =[ ]
  FOR s=1 to Patients_Symptoms _Count
  Patients_Symptoms [i]=Patients_Symptoms [i]&_
Patients_Symptoms
  NEXT s
  Return All
End Sub

G. ALGORITHMS TO BUILD THE INFERENCE ENGINE

Step 1: loop through all Weighted Disease's Symptoms List
{
Step 2: compare Patient's Symptoms List with disease symptoms' list
Step 3: if match found then return Disease Name
Step 4: if Match is not found then return all diseases' name whose symptoms Set intersect with patient's symptoms Set and the degree of such intersection in decreasing order
Step 5: select the disease name with the highest degree of intersection in [3] as the diagnoses disease and refer for confirmation test in phase II
Step 6: if confirmation or output in phase II returns a different output in step 5 then select the next higher degree as generated in step [4] and repeat step [5]
Step 7: repeat step [5] until the confirmation is found or all intersecting sets are exhausted
}

H. PSEUDOCODE FOR INFERENCE ENGINE

```
Sub Inference_Engine
  FOR k=1 to Disease_Count
    IF Patients_Symptom_Set (Intersects_With)
    Disease_SymptomsSet THEN
      Disease(k)=Disease & Degree_of_Intersection
ELSE
NEXT k
RETURN (All Disease(k))

FOR m= maximum(Degree_of_Intersection) TO 1 STEP -1
Output_in_phase I =Disease(m)
IF Output_in_phase I =Output in Phase II then
RETURN (Output in Phase II as Confirmed_Disease)
Exit FOR
ELSE
END IF
NEXT m
```
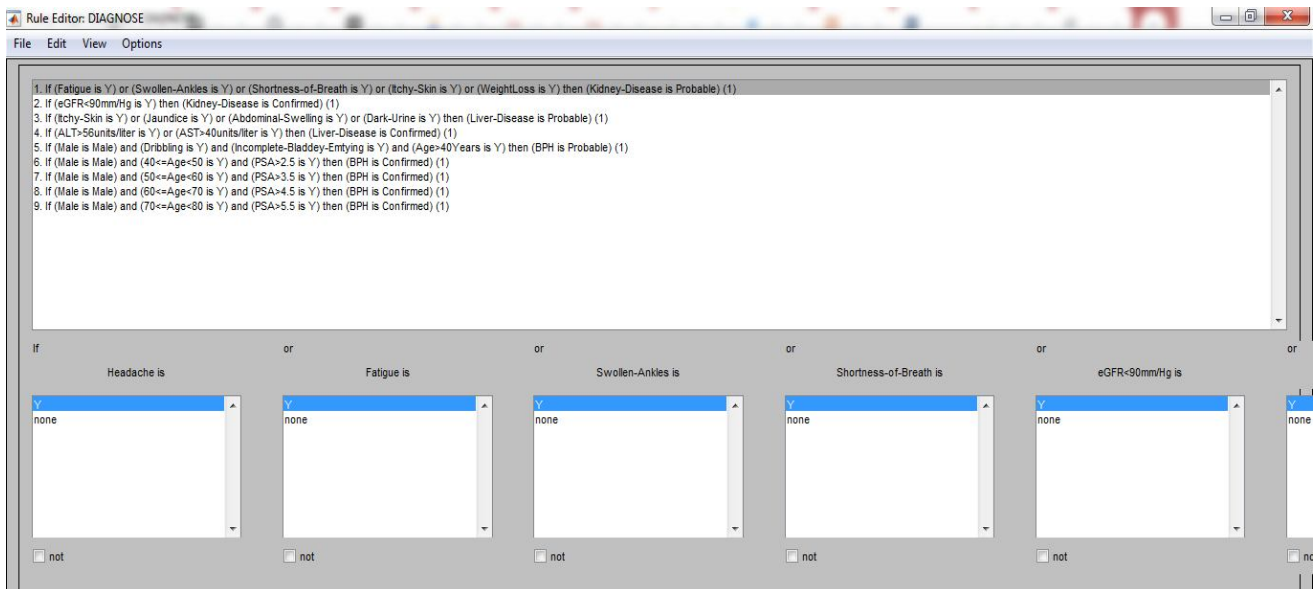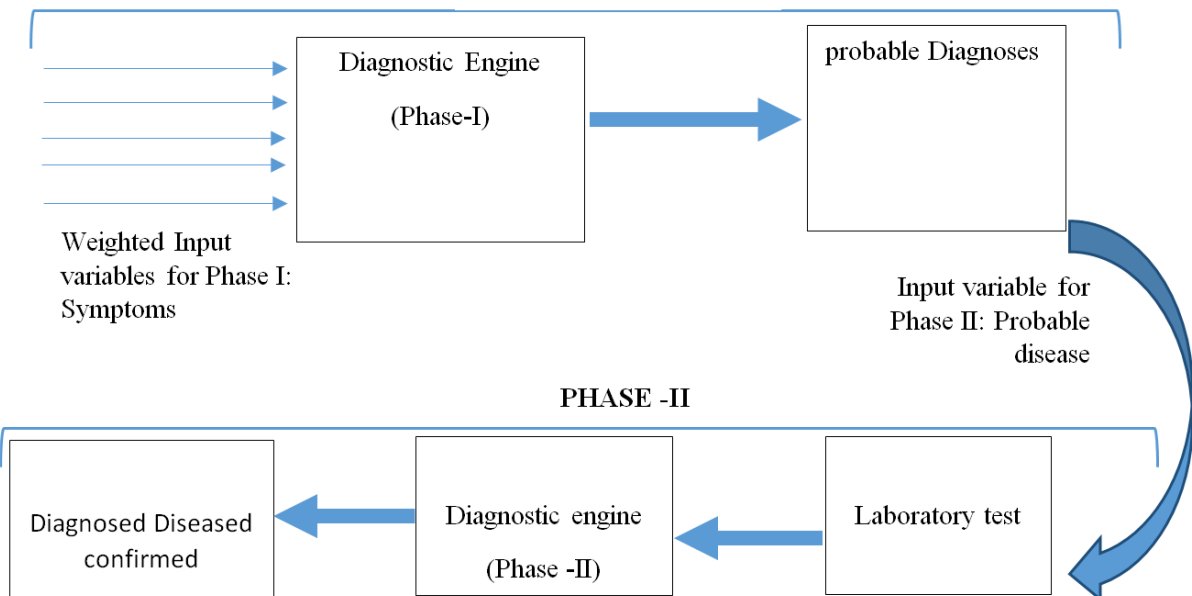

Figure 4. *Fuzzy rules editor.*

Figure 5. *Multiphase Medical diagnostic engine.*

5. RESULTS AND DISCUSSION

Table 1. Prediction results without using the system.

1st Prediction Results				2nd Prediction Results				3rd Prediction Results				4th Prediction Results			
TP	35	FP	29	TP	40	FP	17	TP	26	FP	31	TP	43	FP	11
FN	15	TN	21	FN	10	TN	33	FN	24	TN	19	FN	07	TN	39
Σ	50		50	Σ	50		50	Σ	50		50	Σ	50		50
TPR	0.70	FPR	0.58	TPR	0.80	FPR	0.34	TPR	0.52	FPR	0.62	TPR	0.86	FPR	0.22
		TNR	0.42			TNR	0.66			TNR	0.38				0.78

Average TPR= ΣTPR/4=2.88/4=0.72 ie 72%
Average FPR= ΣFPR/4=1.76/4=0.44 ie 44%
Average TNR= ΣTNR/4=2.24/=0.56 ie 56%

Table 2. Prediction results using the developed system.

1st Prediction Results				2nd Prediction Results				3rd Prediction Results				4th Prediction Results			
TP	44	FP	8	TP	37	FP	20	TP	47	FP	2	TP	45	FP	7
FN	6	TN	42	FN	13	TN	30	FN	3	TN	48	FN	5	TN	43
Σ	50		50	Σ	50		50	Σ	50		50	Σ	50		50
TPR	0.88	FPR	0.16	TPR	0.74	FPR	0.40	TPR	0.94	FPR	0.04	TPR	0.90	FPR	0.14
		TNR	0.84			TNR	0.60			TNR	0.96			TNR	0.86

Average TPR= ΣTPR/4=3.46/4=0.87 ie 87%
Average FPR= ΣFPR/4=0.74/4=0.19 ie 19%
Average TNR= ΣTNR/4=3.26/=0.82 ie 86%

Table 3. Summary of the diagnostic system evaluation.

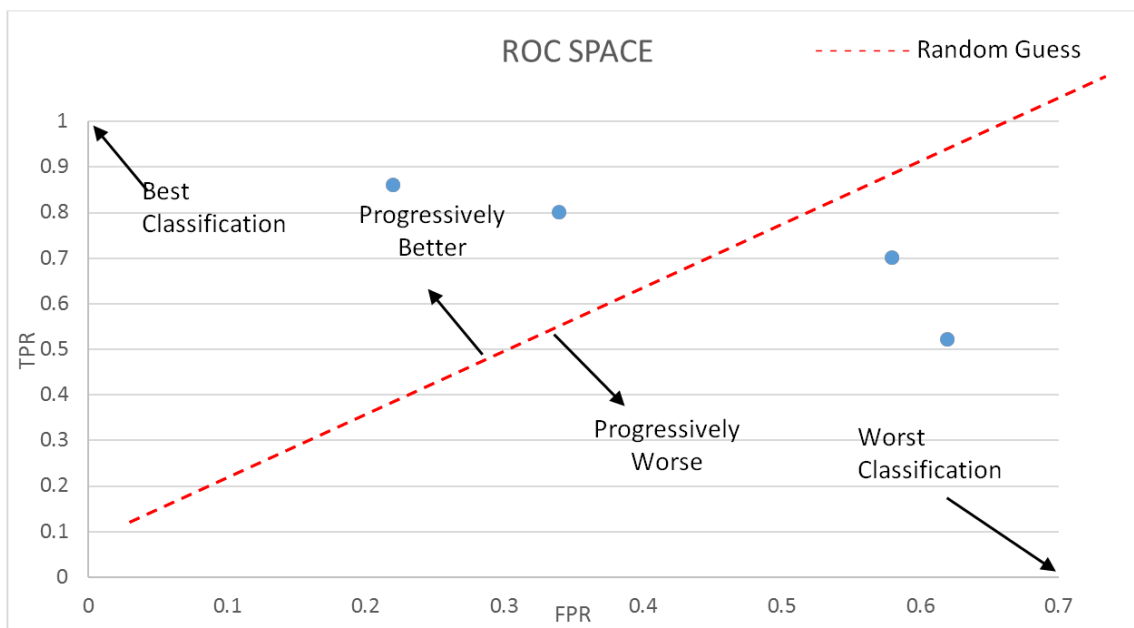| PARAMETER	WITHOUT USING THE SYSTEM	USING THE SYSTEM
TPR	72%	87%
FPR	44%	19%
TNR	56%	86%

Figure 6. Graphical Illustration of the ROC Space and plots of prediction points without using the system.

In order to evaluate the system, four sets of diagnoses data were obtained where the system was not used and then the system was used to see if the diagnoses would have been different using 50 patients' data. The true positive (TP), false Positive (FP) and the true negative(TN) statistics were obtained from where the true positive rate (TPR) or sensitivity, False positive rate (FPR) and true negative rate (TNR) were derived. TPR or Sensitivity is the ability of a test to correctly identify those with the disease and TNR or specificity is the ability of the test to correctly identify those without the disease. The ROC analysis is done for both scenario, using the system and without using the system The ROC curves are as seen in Figure 6 and Figure 7. TPR, FP and TNR are computed using (2) and (3).

$$TPR = \frac{\Sigma\ True\ positive}{\Sigma Condition\ Positive} = \frac{\Sigma\ TP}{\Sigma P} = \frac{\Sigma\ TP}{\Sigma(TP+FN)} = 1-FNR \qquad (2)$$

$$FPR = \frac{\Sigma\ False\ Positive}{\Sigma Condition\ Negative} = \frac{\Sigma\ FP}{\Sigma N} = \frac{\Sigma\ FP}{\Sigma(FP+TN)} \qquad (3)$$
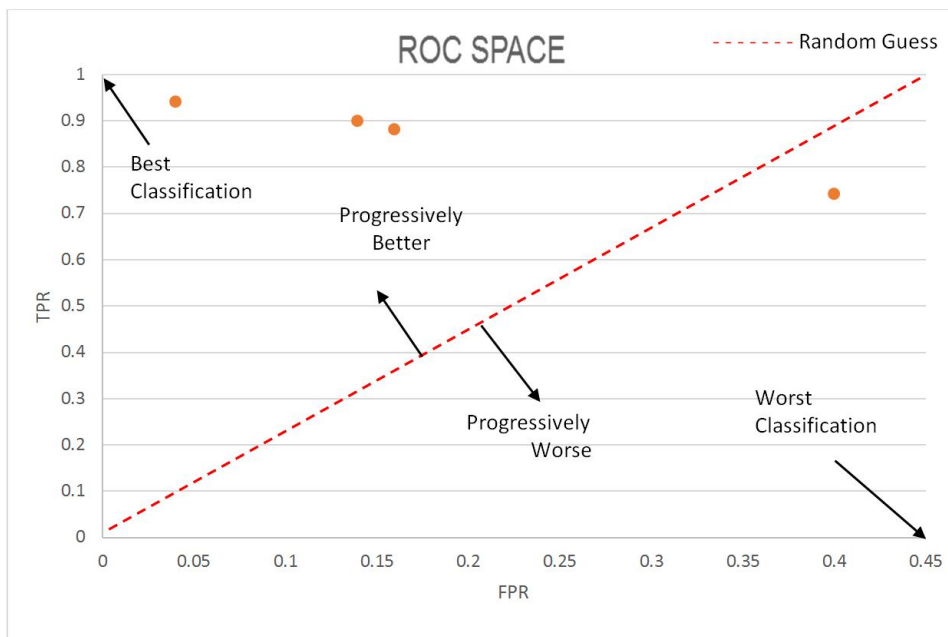
Figure 7. Graphical Illustration of the ROC Space and plots of prediction points using the system.

results of the evaluation.

Where TP means True Positive, FN means False Negative, FP means False Positive, TN means True Negative True positive rate (TPR) and is also referred to as Recall, Sensitivity, probability of detection or Power FNR is False Negative Rate, FPR means False positive rate, Fall-out or the probability of false alarm. Here the graph of TPR vs FPR was plotted from where the quality of the diagnoses could be derived from the receiver operating Characteristics (ROC) space and the results are as presented in evaluation are seen in Table 1 and Table 2. Table 1 shows the results without using the system while Table 2 shows the result when the system is used while table 3 shows the summary of the

The result shows that sensitivity which is the ability of a test to correctly identify those with the disease (True Positive Rate) was and specificity which is the ability of the test to correctly identify those without the disease also called True Negative Rate (TNR) stood were 87% and 86% respectively and by using the system, they stood at 72% and 56% respectively. The results also show that the false positive rate (FPR) which indicates the degree of false alarm is 19% using the new system while it is 44% without using the system. Figure 6, Figure 7 and Figure 8 are all illustrations of the results of the evaluation.
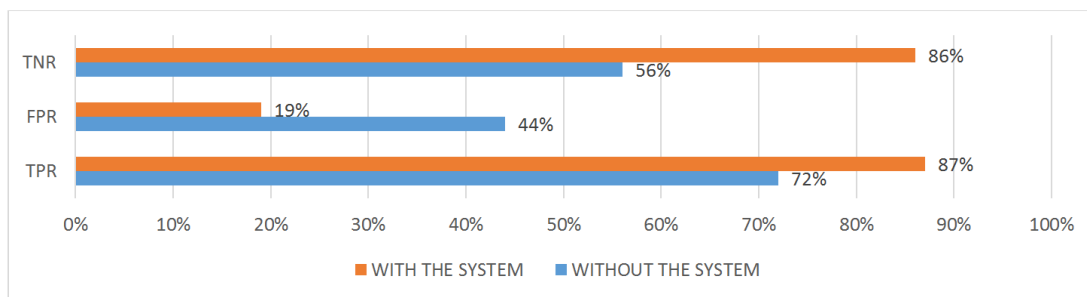

Figure 8. Chart illustrating the evaluations of using clinical decision support system (CDSS) over Non usage of CDSS

6. CONTRIBUTION TO KNOWLEDGE

A good model to minimize diagnostic errors using weighted input variables and rule based fuzzy logic in clinical decision support systems has been developed and

Efficient algorithms that use weighting to properly map symptoms to the appropriate disease including diseases with common symptom sets has also been developed.

7. FUTURE WORK

In the future, the model's knowledge base and the inference engine could be extended to cover diseases that could be diagnosed using imaging.

8. CONCLUSIONS

The mortality rate from diseases due to delayed, imprecise or outrightly wrong diagnoses is alarming and is worsened by low level of literacy, poverty, ignorance, lack of adequate and well trained medical personnel especially in low income countries. Since most of the diseases have common symptoms, diagnosing them correctly could pose a major challenge to health personnel. Considering the necessity to find a better solution, a model was developed in this research to minimize diagnostic errors and improve on the sensitivity, specificity and accuracy of disease diagnoses using fuzzy logic rules and weighted input variables which ensured the right contribution of each symptom to the whole classification process and enabled a proper mapping of inputs(Symptoms) to result(diseases). Evaluation the model using different parameters shows that the aim has been met which ensures that clinical errors could be significantly reduced implying that patients are not taken towards the wrong management route, diagnoses are not delayed and the wrong treatment are not administered on patients

REFERENCES

[1] Ahmed A. E, Sherif E., & Ahmed E. (2011). A fuzzy decision support system for management of breast cancer. *International Journal of Advanced Computer Science and Applications* 2(3):34-40.

[2] Das S., Guha, D. & Dutta B. (2016). Medical diagnosis with the aid of using fuzzy logic and intuitionistic fuzzy logic. *The International Journal of Research on Intelligent Systems for Real Life Complex Problems* 45:850–867.

[3] Graber M. L., Franklin N. & Gordon R. (2005). Diagnostic error in internal medicine. *Archives of Internal Medicine* 165(13): 1493-1499.

[4] Hossein G., Omid P., Javad H., Gholamali R. & Mohammad M. (2011). Diagnosing breast cancer with the aid of fuzzy logic based on data mining of a genetic algorithm in infrared images. *Middle East Journal of Cancer* 3 (4): 119-129.

[5] Johnston M., Brian R. & Haynes A. (1994). Effects of computer-based clinical decision support systems on clinician performance and patient outcome: a critical appraisal of research published: *Annals of Medicine* 120 (2): 135-142. DOI: 10.7326/0003-4819-120-2-199401150-0000.

[6] José A., Mikel G., Aranzazu J., Antonio B. Miguel P., and Humberto B. (2014). Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system. *Applied Soft Computing* 20:103-111.

[7] Manish R. and Sedamkar R. (2013). Design of expert system for medical diagnosis using fuzzy logic. *International Journal of Scientific& Engineering Research* 4 (6): 2914-2921.

[8] Muhammad Q., Saleem A., Yi L. & Muhammad N. (2020). Multi-criteria decision support systems based on linguistic intuitionistic cubic fuzzy aggregation operators. *Journal of Ambient Intelligence and Humanized Computing. DOI*: https://doi.org/10.1007/s12652-020-02563-1

[9] Preety Dagar, Aman J. & Deepti G (2015). Medical diagnosis system using fuzzy logic toolbox. *International Conference on Computing, Communication and Automation* (ICCCA2015). DOI: 10.1109/CCAA.2015.7148370.

[10] Sanchez E. (1996). Truth-qualification and fuzzy relations in natural languages application to medical diagnosis. *Fuzzy Sets and Systems* 84(2).155-167.

[11] World Health Organization (2016). Diagnostic errors: technical series on safer primary care. Geneva: Licence: CC BY-NC-SA 3.0 IGO.

[12] World Health Organization (2020). Global health observatory data repository: number of new HIV infections estimates by WHO region https://apps.who.int/gho/data/view.main. HIVINCIDENCEREGIONv?lang=en. Online; accessed 19-May-2020]

[13] World Health Organization (2016). Partners urge action to tackle HIV misdiagnosis. https://www.who.int/hiv/mediacentre/news/hiv-misdiagnosis/en/. [Online; accessed 19-May-2020]

[14] Yang Z., Li X., Garg H., Qi M.(2020). Decision Support Algorithm for selecting an antivirus mask over COVID-19 pandemic under spherical normal fuzzy environment. International Journal of *Environmental Research and Public Health 17(10):*3407. DOI:https://doi.org/10.3390/ijerph17103407

[15] Giovanni I., Valeria M., Donatella V., Stefania S., Maria T. (2020). Fuzzy logic–based clinical decision support system for the evaluation of renal function in post transplant patients. *Journal of Evaluation in Clinical Practice 26(4) :*1224-123