MAL-TYPHO DIAGNOSIS INTELLIGENT SYSTEM (MATDIS): THE AUTO-DIAGNOSTIC RULE GENERATION ALGORITHM.

S.C. Chiemeke
University of Benin
Benin City
Edo State,. Nigeria
schiemeke@yahoo.com

E.U. Omede
Department of Computer Science
Delta State University
Abraka, Delta State, Nigeria

ABSTRACT
With advancement in medical knowledge systems, determining useful knowledge is becoming more difficult as the use of conventional techniques are no more effective in diagnosis, hence the aid of Intelligent Systems are becoming inevitable in disease diagnosis. One major challenge in designing any intelligent system is the cost of extracting knowledge from human expert for a constructive knowledge base. The difficulty is due to factors such as unavailability of sufficient experts, unwillingness of some experts to share knowledge, human limitations in knowledge acquisition. To proffer solution to this problem, the proposed system provided an algorithm that automatically generates set of diagnostic rules for diseases (Malaria and Typhoid fever is considered in this study) diagnosis using Symptoms Manifestation Degree (SMD) which were optimized with genetic algorithm. The algorithm was simulated using MatLab2013R, and the generated rules were evaluated using statistical F-test analysis computation to compare them with rules from different human experts. F-test returns the values 0.9713 and 0.6101 for Malaria and Typhoid diagnostic rules respectively proving that no significance difference exists between them.

Keywords*:* Intelligent System, Auto-diagnostic rule algorithm, Genetic algorithm

1. INTRODUCTION

Malaria and Typhoid fever are among the life threatening febrile diseases in developing countries. They are though curable but are complicated when neglected and not properly diagnosed and treated. These diseases, though caused by different micro-organism Plasmodium and Salmonella typhi respectively, are often present with mimicking symptoms especially in the early stages of typhoid fever (Ammah et al. 1999, Ohanu et al. 2003). This situation often presents diagnostic problem and in some cases could lead to diagnostic confusion (Uneke, 2008). Most common convectional diagnosis methods for malaria are Clinical, Blood smear while for Typhoid fever are Clinical, Widal test and blood culture. These methods are highly demanding as they require skillfulness, experts and laboratory equipment availability. The need for proper and prompt diagnosis Malaria and Typhoid fever for their effective management has provoked advances in development of Expert or Cognitive systems since emergence of Information Technology which opened the opportunity. Extraction of sufficient and accurate knowledge from human experts remains one of the major challenges in this new innovation. Hoffman (1987) notes that extraction and characterization of knowledge and skills of an expert is a major bottleneck in system development process and the extraction process can be laborious (Quinlan, 1984).

Malaria and Typhoid fever often manifest different symptoms in varied victims; this has caused variations diagnostic conditions or set of rules in systems designed for Malaria or Typhoid fever diagnosis. The literatures reviewed reveals inconsistency and insufficient diagnostic rules for proper and extensive system training. System can be said to be intelligent only when it can learn and be able to adapt to future changes, and more extensive the training data set, the more adaptive the system. It is in this regard that the proposed system (MATDIS) provides an algorithm for generation sufficient set of rules for extensive training of the system.

2. RELATED WORKS

In Oguntimilehin et al. (2013) a typhoid clinical diagnosis system was developed using machine learning approach. The learning technique was used on labeled set of typhoid fever conditional variables to generate rules for diagnosis of typhoid fever. Samuel and Omisore (2013) proposed a hybrid intelligent system for diagnosis of typhoid fever, the system was modeled with neuro-fuzzy technique, and the rule base for fuzzy inference constituted of the rules formulated experts in the medical field. Agakar and Ghatol (2010) used artificial intelligent algorithm for diagnostic-cum preventing approach to immaturity, fragility and vulnerability in infants.

The system was designed to manage disease in infants and classify detected disease into either Malaria, Typhoid fever or Dengue. Adehor and Burell (2008) designed n expert system to combine the action-oriented Integrated Management of Child illness (IMCI) and disease-oriented Health Information System (HIS). The system is used to manage problem of classification in IMCI-HIS system whose difference classification approach was used to construct the rule base and its classification is based on certainty factors. Adekoya et al. (2008) proposed Medical Expert System (MES) for diagnosis and therapy recommendation for managing tropical diseases. Donfack-Kana et al. (2009) proposed Online System for Diagnosis and Treatment of Malaria (OSFDATOM) is a web-based expert system designed for diagnosis of Malaria. Djami et al. (2011) presented fuzzy expert system for the management of malaria (FESMM) for providing decision support platform to malaria researchers, physicians and other healthcare practitioners in malaria endemic regions. Adetunmbi et al. (2012) used machine learning technique, rough set on labeled sets of malaria fever symptoms collected to generate explainable rules for each level of severity and appropriate therapy is provided. Olabiyisi et al. (2011), designed a system to diagnose tropical diseases which include Malaria and Typhoid among others.

Prihatini and Putra (2012) combined fuzzy logic and certainty factor to design system for diagnosis of tropical infections. From the reviewed works, it was observed that series of systems have been designed for diagnosis of either Malaria or Typhoid fever or both using different techniques from convectional through machine learning to soft computing. However each case rely solely on experts knowledge extract for rule base construction, and there is disparities in set of rule generated for a particular disease as is revealed in the literature, this brought to mind that none of the set is sufficient in its self and may not have exhausted all possible conditions. As a solution, this research which employs hybridized technique in designing intelligent system for diagnosis of Malaria and Typhoid fever incorporated an auto-rule generation algorithm which maps the determinant features with infestation period to generate expansive set of rules for training of proposed system.

3. MATERIALS AND METHODS

The input (optimized symptoms) to the system are output of genetic optimization of data (the attributes for diagnosis of malaria and typhoid fever and classification criteria) gotten from series of consultations with medical experts (especially those in Delta State University's medical centre) and standard literatures in tropical medicine field while the manifestation degree of each symptom was obtained from expert records of infected patients.

Table 1: shows different categories of symptoms of both malaria and typhoid fever with their manifestations in infected patient at different periods from the point of infestation. M and T represents Malaria and Typhoid fever respectively, "a" means less or equal to 2 weeks of infestation while "b" means greater than 2 weeks. Time factor (Period) is very important because most of the symptoms are time dependent. "*" indicates the presence of a symptom.

The manifestation degree of symptoms at different period is indicated as:
A ≡ "Almost all"; V ≡ "Very common"; C ≡ "Common"; U ≡ "Uncommon"; R ≡ "Rare"
Vr ≡ "Very rare"; Xr ≡ "Extremely rare"; "-"≡ Symptom not associated with the disease.

Table 1: Categorized Symptoms of Malaria and Typhoid Fever with degree of periodic manifestation. (Medscape Clinical Presentation, 2012).

S/No	Category	Symptom	Code	M	T	a	b
1	Systemic Features (X_1)	High fever	S_1	*		Vc	
2		Stepwise Fever	S_2		*	Vc	Vc
3		Chills	S_3	*	*	U/A	-
4		Diaphoresis(Excessive sweating)	S_4	*	*	Vc	-
5		Rigors (exaggerated chill)	S_5	*	*	A/U	Uc
6		Anorexia (Loss of appetite)	S_6	*	*	C/A	-
7		Lethargy (Fatigue)	S_7	*		A	-
8		Insomnia	S_8		*	Vc	-
		Weight loss	S_9	*	*	C/C	
9	Neurologic Features(X_2)	Malaise(ill feeling)	S_{10}	*	*	C/A	C
10		Frontal Headache	S_{11}		*	Vc	Vc
11		Headache	S_{12}	*		C	C
12		Psychosis (mental disability)	S_{13}		*	Vr	C
13		Confusion /delirium	S_{14}	*	*	R/Vc	-
14	Pulmonary Feature(X_3)	Bronchitic Cough	S_{15}		*	C	
15		Cough	S_{16}	*		A	A
16		Rales (sound from unhealthy lung)	S_{17}		*	C	-
17		Mild cough	S_{18}		*	C	-
18		Pneumonia	S_{19}		*	R	C
19	Ear, Nose , Throat Feature (X_4)	Coated Tongue	S_{20}		*	Vc	-
20		Epistaxis(Nose bleed)	S_{21}	*		R	
21	Dermatologic Feature (X_5)	Rose spot	S_{22}		*	R	-
22	Cardiovascular Feature (X_6)	Dicrotic pulse	S_{23}		*	R	C
23		Thrombophlebitis (blood clot)	S_{24}		*		Vr
24	Gastro intestine Feature(X_7)	Nausea and Vomiting	S_{25}	*		C	C
25		Diarrhea	S_{26}	*	*	C/-	C/C
26		Jaundice	S_{27}	*	*	C/C	C/-
27		Constipation	S_{28}		*	Vc	C
28		Bloating	S_{29}		*	Vc	-
29		Intestine hemorrhage	S_{30}		*	Vr	Vc
30		Splenomegaly	S_{31}	*		-	R
31		Hepatosplenomegaly	S_{32}		*	C	
32		Intestinal Perforation	S_{33}		*	-	R
33	Musculoskeletal Feature(X_8)	Myalgia (Muscle pain)	S_{34}	*		Vc	
35		Arthralgia (Joint pain)	S_{35}	*		Vc	
36	Urogenital Feature (X_9)	Urinary retention	S_{36}		*	C	
37		Renal pain	S_{37}	*	*	-/R	Vr/-

3.1 Genetic Algorithm (GA) For Diagnostic Features Selection

GA is an adaptive heuristic search algorithm which computational model is inspired by process of natural evolution. It is invented by John Holland in 1960's and developed in collaboration with his students and colleagues at the university of Michigan in the 1960's and 1970's (Mitchell,1988, in Darzi et al. 2011). Detail discussion on genetic algorithm is not within the scope of this paper. The concern of this work is to employ GA for selection of determinant features for generation of diagnostic rules for Malaria and Typhoid fever. The GA parameter setup used for features selection is given below:

Search Method = GA
Crossover function = single point
Mutation rate (Uniform) = 0.01
Population Size = 10
Generation= 50

Evaluation function, $F_D = D_e - D_o$ ..... (1)

Where D_e is the Expected Manifestation degree (how much globally a symptom should manifest to be factor for ascertaining presence of a disease, in this case referred as "Almost all" rated as 5).
D_o is the Observed Manifestation degree (actual occurrence for a symptom as observed in sampled infected patients).

The fitness function, $F = \frac{1}{1+F_D}$ ..... (2)

F_D is the evaluation function (also known as objective function) which provides the measure of the relevance of a symptom in a chromosome set for a disease diagnosis and classification. Fitness function (F) transforms this measure of relevance into an allocation of reproductive opportunities by measuring the value of the individual in relation to the rest of the population. It is given by equation (5). Maximum optimization method is applied for computing the fitness of each chromosome.

For selection of parents for crossover based on their fitness, roulette –wheel selection is used.

The probability (p) of k^{th} string in the population being selected is given by

$p_k = F_k / (\sum_{i=1}^{m} F_i)$ ....(3)

That is the ratio of a strings fitness to the sum of the fitness of all the strings in the population;

F_k is fitness for string k in the population, m is the number of individuals in the initial population.

Table 2: Summarized Evaluation of initial population

String No	Initial population	D value	Fitness(F_k) $F = \frac{1}{1+F_D}$	P_k	Expected Count P_k*m
1	011	3	0.33	0.11	1.1
2	001	1	0.20	0.07	0.7
3	010	2	0.25	0.09	0.9
4	001	1	0.20	0.07	0.7
5	000	0	0.16	0.05	0.5
6	011	3	0.33	0.11	1.1
7	100	4	0.50	0.17	1.7
8	010	2	0.25	0.09	0.9
9	001	1	0.20	0.07	0.7
10	100	4	0.5	0.17	1.7
Sum			2.92	1	10
Average			0.29	0.13	1.3
Max			0.50	0.17	1.7

The determinant symptoms are selected at fitness of 0.5 that is those with 65% and above manifestation degree.

Table 3: Output from GA optimization of the symptoms

S	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}	S_{17}	…	S_{33}	S_{34}
D	4	5	4	2	5	5	4	3	5	4	3	1	4	3	3	3	1	…	3	1
Fs	0.5	1.0	0.5	0.25	1.0	1.0	0.5	0.33	1.0	0.5	0.33	0.2	0.5	0.33	0.33	0.33	0.2	…	0.33	0.2
Fd	0	0.5	0	-0.25	0.5	0.5	0	-0.17	0.5	0	-0.17	-0.3	0	-0.17	-0.17	-0.17	-0.3	…	-0.17	-0.3

The function considered 34 symptoms as shown in Table 3 and filtered 23 significant symptoms, that is those with $F_d >= 0$. d is the values of the manifestation degree of the symptoms. F_s is individual fitness ranking and F_d is deviation of F_s from evaluated fitness function value. The determinant symptoms which are used as input to auto rules generation algorithm are shown in Table 4

Table 4: The Predominant symptoms for each period

Disease	$\leq 2\,weeks$	$> 2\,weeks$
Malaria	S_1	
	S_4	
	S_5	S_{16}
	S_7	
	S_{26}	
	S_{27}	
	S_{34}	
	S_{35}	
Typhoid Fever	S_3	S_2
	S_4	S_{11}
	S_6	
	S_8	
	S_{10}	S_{26}
	S_{14}	S_{28}
	S_{20}	S_{30}
	S_{29}	
	S_{33}	

Table 4 shows the number of predominant symptoms of each symptom for Malaria and Typhoid fever in less than two weeks of infestation and more than two weeks of infestation. The predominant symptoms with Manifestation degree "Almost all" or "Very common" are considered as determinants for auto rule generation.

3.2 Determination of the possible diagnostic rules from symptoms categories.

Three degrees of severity; Mild, Moderate and Severe are considered in this work and are represented with i,j,k respectively for each symptom.

Let s and t, be two different symptoms, and i, j, k, be possible degrees of severity each symptom can assume, then considering the possible mappings the number of rules that can be generated is given by:

$$Set\ of\ rules = \begin{Bmatrix} s_i t_i\ ;\ s_i t_j ; s_i t_k \\ s_j t_i\ ;\ s_j t_j ; s_j t_k \\ s_k t_i\ ;\ s_k t_j ; s_k t_k \end{Bmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

with 2 symptoms, there exist 9 possible rules.

Let consider x, y, z as a set of symptoms and i,j,k as 3 possible degrees of severity for each symptom, then the possible number of rules are:

$$Set\ of\ rules = \left\{ \begin{matrix} x_i y_i z_i ; x_j y_i z_i ; x_k y_i z_i \\ x_i y_i z_j ; x_j y_i z_j ; x_k y_i z_j \\ x_i y_i z_k ; x_j y_i z_k ; x_k y_i z_k \\ x_i y_j z_i ; x_j y_j z_i ; x_k y_j z_i \\ x_i y_j z_j ; x_j y_j z_j ; x_k y_j z_j \\ x_i y_j z_k ; x_j y_j z_k ; x_k y_j z_k \\ x_i y_k z_i ; x_j y_k z_i ; x_k y_k z_i \\ x_i y_k z_j ; x_j y_k z_j ; x_k y_k z_j \\ x_i y_k z_k ; x_j y_k z_k ; x_k y_k z_k \end{matrix} \right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (5)$$

From Equation 5, (each cell representing a rule) there exist 9x3 matrix; resulting to total number of 27rules.
This implies that for X number of degrees of Severity and Y number of symptoms;

$$Rule\ Matirx\ function, R_{xy} = (X^{(y-1)}, X) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (6)$$

Algorithm for mapping the symptoms and the degree of severity to form the rules.
Declare as integer, X, Y, s,i, r; and s as arrays
Initialize X,s,i to 0
Input X,s,i
Loop Y=X to 1 step -1
    Loop s = 1 to i
        R ((X^i)-1) = $Y_1(1), Y_2(1) \ldots Y_n(i)$ …………………………………… (7)
End loop
End loop

Figure 1: Auto rules generation Algorithm.

Applying this principle in our auto rules generation algorithm, different categories of symptoms in malaria and typhoid fever will present the possible number of rules formation as is shown in Table 5 below:

Table 5: Possible number of rules that can be generated

Category	Malaria		Typhoid fever	
	<=2wks	>2wks	<=2wks	>2wks
	No. of Symptom	No. of Symptom	No. of Symptom	No. of Symptom
Systemic	4	-	4	1
Neurologic	-	-	4	1
Pulmonary	1	1	-	-
Ear, Nose, Throat	-	-	1	-
Gastro intestine	-	-	2	1
Musculoskeletal	2	-	-	-
Possible number of rules	3^7	3^1	3^{11}	3^3

3.3 Experiment and Result
The auto-rule generator algorithm was simulated using Matlab R2013a simulator, and 19630 rules were generated for diagnosis of Malaria and Typhoid. The sample is shown in Figure 2.

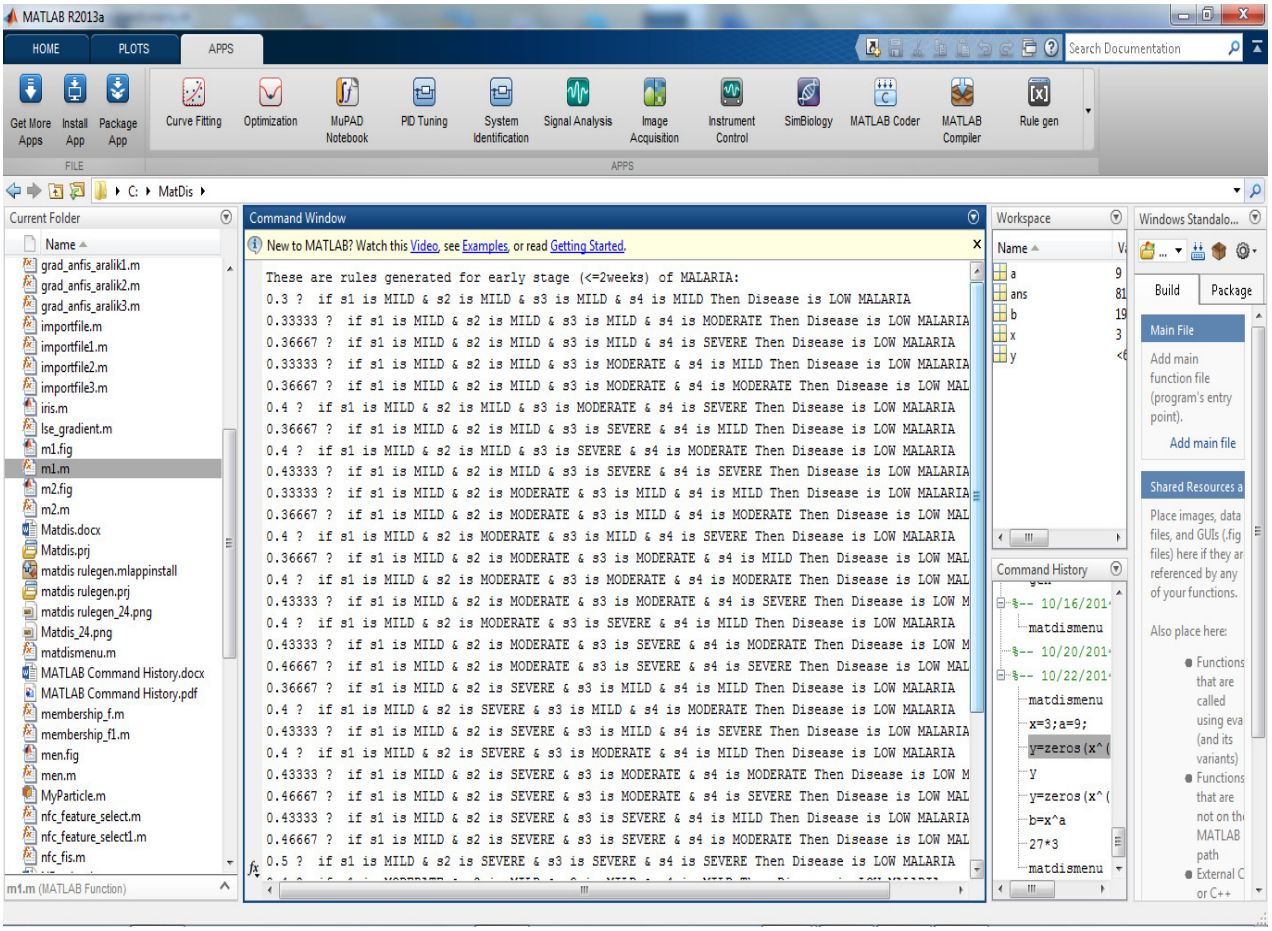

Figure 2: Sample of auto generated rules

3.3.1 Validation of Generated Rule

The generated rules assumed 3 degrees of severity (Mild, Moderate and Severe) for symptoms and two (Low and High) for diseases. These are assigned the real values Mild=0.3, Moderate = 0.6, Severe = 0.9. For conclusion values ≤ 0.6 and > 0.6 indicate low and high (Malaria and Typhoid fever) respectively. The use of real values is for easy mathematical calculation. The generated rules were validated by assigning the same values to the symptoms severity, only the values of determinant symptoms were used to evaluate the conclusion by calculating their average. Out of 200 and 50 samples of diagnostic rules for Malaria and Typhoid fever respectively formed by the domain experts shown in Table 6 and Table7, conclusion of 185 of the rules for Malaria and 43 of Typhoid fever correspond to conclusion of the auto generated rules. The accuracy percentage is given by:
Accuracy% = (Matched/ (Matched + Unmatched))* 100

For Malaria rules = (185/ (185+15))*100
= 92.50%

For Typhoid fever rules = (43/ (43+7)*100
= 86%

Table 6: Malaria Diagnostic Rules From Experts.

Rule No	S1	S12	S25	S27	S31	S35	S7	S34	S14	S6	MP		X-Conclusion	A-Conclusion	Evaluation of X-Conclusion	Evaluation of A-Conclusion
1	0.3	0.3	0.3	0.3	0.3	0.6	0.3	0.3	0.3	0.6	0.3	0.3	0.3	0.45	1	1
2	0.6	0.3	0.3	0.3	0.3	0.3	0.6	0.6	0.3	0.9	0.9	0.6	0.6	0.525	1	1
3	0.9	0.6	0.3	0.3	0.3	0.3	0.6	0.9	0.6	0.9	0.9	0.6	0.9	0.675	2	2
4	0.9	0.3	0.3	0.3	0.3	0.3	0.9	0.9	0.6	0.3	0.3	0.9	0.9	0.75	2	2
5	0.6	0.3	0.3	0.6	0.3	0.3	0.3	0.6	0.6	0.6	0.9	0.6	0.6	0.6	1	1
6	0.3	0.6	0.6	0.3	0.3	0.3	0.3	0.6	0.9	0.6	0.3	0.3	0.3	0.525	1	1
7	0.3	0.3	0.6	0.6	0.3	0.3	0.9	0.9	0.9	0.6	0.6	0.6	0.9	0.75	2	2
8	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.9	0.6	0.6	0.6	0.6	0.825	1	2
9	0.6	0.3	0.3	0.3	0.3	0.9	0.3	0.6	0.3	0.9	0.6	0.6	0.6	0.6	1	1
10	0.3	0.3	0.6	0.6	0.3	0.6	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.525	1	1
11	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.3	0.9	0.9	0.6	0.9	0.975	2	2
12	0.6	0.9	0.6	0.9	0.6	0.9	0.6	0.9	0.6	0.6	0.3	0.6	0.9	0.9	2	2
13	0.3	0.6	0.6	0.6	0.3	0.3	0.3	0.6	0.6	0.3	0.6	0.6	0.6	0.525	1	1
14	0.9	0.9	0.6	0.9	0.9	0.9	0.9	0.9	0.3	0.6	0.6	0.9	0.9	0.975	2	2
15	0.3	0.3	0.3	0.6	0.9	0.3	0.3	0.9	0.9	0.3	0.9	0.6	0.9	0.6	2	1
16	0.9	0.6	0.9	0.6	0.9	0.6	0.9	0.9	0.6	0.9	0.3	0.6	0.9	0.9	2	2
17	0.3	0.9	0.6	0.6	0.3	0.6	0.3	0.6	0.3	0.9	0.3	0.6	0.9	0.525	2	1
18	0.6	0.9	0.9	0.3	0.9	0.9	0.6	0.9	0.9	0.9	0.9	0.9	0.9	0.825	2	2
19	0.6	0.6	0.6	0.6	0.3	0.3	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.675	1	2
20	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.3	0.9	0.9	0.6	0.9	0.975	2	2
21	0.9	0.9	0.9	0.9	0.9	0.3	0.9	0.6	0.6	0.6	0.9	0.9	0.9	0.9	2	2
22	0.6	0.9	0.3	0.3	0.3	0.9	0.6	0.9	0.3	0.3	0.9	0.9	0.9	0.675	2	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	1	1
100	0.3	0.3	0.3	0.6	0.3	0.3	0.6	0.3	0.3	0.6	0.3	0.3	0.3	0.525	1	1
101	0.6	0.9	0.9	0.9	0.6	0.6	0.9	0.6	0.9	0.9	0.3	0.9	0.9	0.975	2	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	1	1
200	0.3	0.3	0.6	0.3	0.6	0.9	0.6	0.9	0.3	0.6	0.3	0.3	0.3	0.6	1	1

Table 7: Typhoid fever diagnostic rules from experts.

Rule No	S2	S6	S9	S10	S12	S14	S16	S20	S21	S22	S26	S28	S29	S30	S34	S36	X-Conclusion	A-Conclusion	Evaluation of X-Conclusion	Evaluation of A-Conclusion
1	0.30	0.30	0.30	0.30	0.30	0.30	0.60	0.30				0.30	0.30	0.30	0.30		0.30	0.30	1	1
2	0.30	0.30	0.30			0.90		0.60				0.30		0.30		0.30	0.30	0.45	1	1
3				0.30	0.30	0.60		0.30			0.30	0.30	0.60	0.30	0.30		0.30	0.40	1	1
4	0.60	0.60	0.30		0.30	0.30		0.30	0.30		0.30	0.90					0.30	0.54	1	1
5	0.30	0.90		0.30	0.30							0.30	0.30	0.60	0.30		0.60	0.45	1	1
6	0.60	0.60	0.30	0.30	0.30	0.30		0.90	0.30			0.30	0.60	0.30	0.30		0.60	0.49	1	1
7	0.30		0.30		0.30	0.30		0.30				0.90	0.60	0.90	0.30	0.60	0.90	0.55	2	1
8	0.30	0.30		0.30		0.30		0.60				0.90	0.30	0.60			0.60	0.45	1	1
9	0.30	0.60		0.30	0.90			0.30		0.30	0.30		0.30				0.60	0.36	1	1
10	0.90	0.30	0.30			0.60	0.60	0.60		0.30		0.90	0.60	0.30	0.60	0.30	0.60	0.60	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	1	1
20	0.60			0.90	0.30	0.60		0.30		0.30		0.60	0.90	0.30	0.30		0.90	0.60	2	1
21	0.60	0.90	0.30	0.60	0.30	0.60	0.30	0.60	0.30	0.60	0.90	0.60	0.60	0.90		0.30	0.90	0.68	2	2
22	0.90	0.30	0.90	0.60	0.60	0.90	0.30	0.90	0.60		0.30	0.60	0.60	0.90			0.90	0.71	2	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	1	1
48	0.60	0.90	0.90	0.60	0.30	0.60	0.30	0.90	0.60		0.90	0.60	0.90	0.60	0.30	0.60	0.90	0.71	2	2
49	0.60	0.90	0.30	0.90	0.30	0.90		0.30		0.60	0.30	0.90	0.60	0.90	0.90	0.30	0.90	0.75	2	2
50	0.30	0.60	0.90	0.30	0.90	0.30		0.30				0.30	0.60	0.30	0.30	0.90	0.30	0.38	1	1

3.3.2 Validation using F-Test analysis

F-test statistical analysis was carried out using X-conclusion and A-conclusion which are experts' and auto-rule generator diagnostic conclusions respectively to ascertain that there is no significance difference between the two. The real values <= 0.6 are assigned integer 1 while values >6 are assigned integer 2 to represent low and high (Malaria or Typhoid fever) respectively.

Table 8: F-test statistical Analysis

	Malaria Diagnostic rules		Typhoid fever diagnostic rules	
Statistical functions	X-conclusion	A-conclusion	X-conclusion	A-conclusion
$Observed\ Sample$	200	200	50	50
$Mean$	1.4667	1.4333	1.3333	1.2222
$Standard\ deviation\ \sigma_x$	0.5074	0.5040	0.4851	0.4278
$Variance\ \sigma^2$	0.2575	0.2540	0.2353	0.18301
$F-test$	0.9713		0.6101	
$Critical\ F\ value\ (lower)te$	0.7568		0.5675	
$Critical\ F\ value\ (Upper)$	1.3214		1.7622	

From the analysis result, the hypothesis that $\sigma^2\ (X-conclusion) = \sigma^2\ (A-conclusion)$ in both cases is accepted because F-test results is located within the critical F values, thus there is no significance difference between Diagnostic rules formed by experts and the ones generated by Auto-rule generator algorithm.

4. CONCLUSION

This paper presented auto-rule generator algorithm design, which is a preprocessor in proposed MatDis system for generation of exhaustive diagnosis rule to enable extensive training of the system in the diseases' domain. Though large number of rules may cause poor convergence rate, this should be taken care of by adapting Modified Polak-Ribiere(MPR) Conjugate Gradient(CG) which is a faster training algorithm preferable for complex system. In our future work, we intend to use these generated rules as input to proposed neuro-fuzzy system for diagnosis and classification of Malaria and Typhoid fever. We intend to implement the model in an environment windows7 or later Operating system, Matlab R2013a and Visual Studio.

REFERENCES

1. Adehor, A.B, and Burell, P.R. (2008) " The Integrated Management of Health Care Strategies and Differential Diagnosis, by Expert System: A Single-Dimension Approach." http://www.waset.org/journals/waset/v44/v44-91.pdf
2. Adekoye, A.F., Akinwale, A.T. and Oke O.E.(2008). A medical expert system for managing tropical diseases. *Proceedings of the Third Conference on Science and National Development.* Retrieved from: www.unaab.edu.ng/journal/index.php/COLNAS/article/.../145/148
3. Adetunmbi A.O, Oguntimilehin A., and Falaki S.O , (2012). "Web-based medical assistant system for malaria diagnosis and therapy", GESJ: Computer Science and Telecommunications, No1(33), pp 42-53. Available online: gesj.internet-academy.org.ge/ download.php?id=1849.pdf&t=1
4. Agarkar A.M. and Ghatol, A.A(2010); "FFANN based Cost Effective Major Infant disease Management" International Journal of Computer Applications (0975 – 8887)Volume 7– No.11. Retrieved from: www.ijcaonline.org/volume7/number11/pxc3871755.pdf
5. Ammah A., Nkujo-Akenji T., Ndip R. and Deas E.(1999) "An update on concurrent malaria and typhoid fever in Cameroon". Trans R Soc Trop Med Hyg; 2: 127–9.
6. Darzi, M., AsgharLiaei, A., Hosseini, M. and Asghari, H. (2011). Feature Selection for Breast Cancer Diagnosis: A Case-Based Wrapper Approach.
7. Djam X.Y, Wajiga, G.M, Kimbi, Y.H and Blamah N.V. (2011). A Fuzzy Expert System for the Management of Malaria*. International Journal of Pure and Applied Sciences and Technology 5(2), pp. 84-108.* ISSN 2229 – 6107. Retrieved from: www.ijopaasat.in
8. Donfack kana A.F., Abdullahi M., Ezugwu A. E and Alkali S.A(2009). Online system for diagnosis and treatment of malaria. Retrieved from: www.abu.edu.ng/publications/2009-06-22-125106_5476.doc
9. Hoffman R. (1987). The Problem of Extracting the Knowledge of Experts from the Perspective of Experimental Psychology. *AI Magazine (8)2* Retrieved from http://aaaipress.org/ojs/index.php/aimagazine/article/viewFile/583/519. 26/09/14
10. Oguntimilehin A., Adetunmbi A.O. and Abiola O.B. (2013). A Machine Learning Approach to Clinical Diagnosis of Typhoid Fever. *International Journal of Computer and Information Technology (ISSN:2279) (2)4*. Retrieved from http://ijcit.com/archives/volume2/issue4/Paper020419.pdf. 26/09/14
11. Ohanu M.E, Mbah A.U, Okonkwo P.O., and Nwagbo F.S.(2003). Interference by malaria in the diagnosis of typhoid using Widal test alone. West Afr J Med; 22: 250–2.
12. Olabiyisi, S.O., Omidiora, E.O., Olaniyan O.M. and Derikoma,O. (2011). A Decision Support System Model For Diagnosing Tropical Diseases Using Fuzzy Logic. *African Journal of Computing &ICT . 4, 2.(2). Pp.1-6 ISSN 2006-1781*
13. Prihatini, P. M. and Putra, I.K.G.D. (2012). Fuzzy Knowledge-based System with Uncertainty for Tropical Infectious Disease Diagnosis. *IJCSI International Journal of Computer Science Issues, 9, (4), 3, ISSN 1694-0814.* Retrieved from: www.IJCSI.org.
14. Quinlan, J. R. 1984 Fundamentals of the Knowledge Engineering Problem In Introductory Readings in Expert Systems, ed. D Michie, 33-46. London: Gordon and Breach.
15. Samuel, O.W. and Omisore, M.O.(2013). Hybrid Intelligent System for the Diagnosis of typhoid Fever. *Journal of Computer Engineering and Information Technology, (2)2.* Retrieved from http://dx.doi.org/10.4172/2324-9307.1000109
16. Uneke, C.J. (2008). Concurrent malaria and typhoid fever in the tropics: the diagnostic challenges and public health implications. *J Vector Borne Dis 45*, pp. 133–142