A MACHINE LEARNING BASED APPROACH FOR THE MANAGEMENT OF TYPHOID AND MALARIA INFECTION

Abisoye Opeyemi. A[1]., Douglas Ibrahim[2], Abisoye Blessing[3], Elisha Richard[3],
[1]Federal University of Technology P.M.B. 65, Minna, Niger State, Nigeria.
[2]Adamawa State College of Agriculture, P.M.B. 2088, Ganye, Adamawa State, Nigeria.
[3]Federal University of Technology P.M.B. 65, Minna, Niger State, Nigeria. [4]Federal Polytechnic, P. M. B. 35, Mubi, Adamawa State, Nigeria.

ABSTRACT

One of the major public health problems are Typhoid and Malaria co-infection, accounting for the death of millions of people every year apart from contributing to economic backwardness. The large number of deaths recorded with malaria and typhoid fever is as a result of many factors includes: Poor diagnosis, self-medication, shortage of medical experts and insufficient hospital medical laboratories. Therefore, the need for an enhanced malaria and typhoid expert system is greatly needed. An Artificial Neural Network machine learning technique was used on the set of malaria and typhoid fever conditional variables to generate explainable rules for. The labeled database was divided into four different levels of severity and classes in Malaria and Typhoid. Out of 14 data that the physician considered as positive, the ANN found that 11 were positive and 3 were negative. Moreover, out of the 11 data that the physician considered negative, the ANN found that 2 were negative and 9 were positive. Therefore, The ANN produces classification accuracy 65.22% accuracy, 57.89% specificity and 100% sensitivity with malaria infection while classification accuracy of 22%, 12% specificity and 100% sensitivity with typhoid infection on both the training set and testing set. Further studies will focus on using different machine learning techniques to handle multiclass infection cases.

Keywords: Machine Learning, Typhoid, Malaria, Co-infection, Neural Network, preprocessing, features, performance, Adamawa State.

1.0 INTRODUCTION

A mosquito-borne infection affecting human being caused by parasitic protozoansknown as *Plasmodium Specie*is known as Malaria [1]. Malaria symptoms typically include headache, fever, vomiting, etc. in very severe cases it can cause yellow skin, coma or death. Symptoms usually begin ten (10) to fifteen(15) days after being bitten[2]. If not adequately treated, people may have occurrences of the disease months later. In those who have recently survived an infection, reinfection usually causes milder symptoms. This partial resistance disappears over months to years if the person has continuing exposure to malaria. Anopheles mosquitoes were discovered as the carriers of this disease and that the parasite is also capable of living partly in man as the secondary host [3].

Out of over a hundred species of plasmodium in existence but four (4) of these species commonly infect human beings [4]. These common species are:
Plasmodium falciparum, that is seen as the major cause of death in Africa, mostly spread after plasmodium vivax and most deadly; plasmodium vivax; the most widely spread species and it causes symptoms that are of mild severity in man; plasmodium malariae – it is capable of persisting in the blood for very long period of time, possibly decades, without even producing symptoms[5].

Since the causes of malaria has been discovered, all efforts have been made by organizations, government and individuals to ensure that this deadly disease is completely eradicated. World Health Organization effortsinclude: distribution of treated nets, anti-malaria drugs, insects repellents and enlightenment to people that reside near river [6].But these challengesare still in existence: insufficient medical practitioners, application of advancement in Information Technology to equalizing the ratio of patients to medical practitioners by developing expert systems that can personify the human experts in the field of diagnosis and therapy of malaria. Consequently, the traditional method of medical diagnostic is plague and imprecision which has cause many life.

Typhoid fever is a bacterial infection that can spread throughout the body, affecting many organs. Typhoid fever is a systematic disease contracted through ingestion of contaminated food or water. It is caused by the bacterium salmonella entricserovarTyphi, which is a pathogen only of humans. The illness may be mild or severe. Paratyphoid is a clinically similar illness (though often less severe), caused by Salmonella enteric serovarParatyphi A, B, or C. these conditions are sometimes referred to collectively as enteric fever [7].

Typhoid fever remains a major public health problem in developing countries of world even in the twenty first century [8]. Unacceptable morbidity and mortality are still recorded in developing countries inspire of the availability of several drugs over the years for the treatment of typhoid fever [9]. The bacterium lives in the ntestine and bloodstream of human. It is spread between individuals by direct contact with the feces of an infected person. If untreated, around 1 in 4 cases of typhoid end in death.
If treatment is given, less than 4 in 100 cases are fatal. *S.typhi* enters through the mouth and spends 1 – 3 weeks in the intestine. After this time, it makes its way through the intestinal wall and into the bloodstream.
From the bloodstream, its spread into other tissues and organs, the immune system of the host can do little to fight back because *S.typhi* can live within the host's cells, safe from the immune system. Typhoid is diagnosed by detecting the presence of *S.typhi* via blood, urine, stool, etc. [10]. The medical field needs appropriate facilities and medical experts from efficient services delivery [11]. Death occur mostly in situations where patients conditions are critical and yet no medical facilities and expert to handle situations [12].Because medications are readily available without a prescription, people with self-diagnosed infections can access treatment without first seeking a formal clinical consultation or laboratory confirmed diagnosis [13].

The advent of computer has led to the development of several algorithms, models and technologies to ensure accuracy and precision and this has greatly reduced the numbers of patients that dead daily in the hospitals and one of such technology is machine learning which is a branch of artificial intelligence. Machine learning algorithms were from the very beginning designed and use to analyse medical sets[14].

Today machine learning provides several indispensable tools for intelligent data analysis. Machine learning technology is currently well suited for analyzing medical data and in particularly there is a lot of work done in medical diagnosis in small specialized diagnostic problems. There are sufficient data available in healthcare sector and tool needed for effective analysis for the discovery of hidden relationship and data trends is deficient [15].

Data about correct diagnosis are often available in the form of medical records in specialized hospitals. Computer-Aided System or Decision Support System (DSS) that can simulate expert human reasoning or service as an assistant of a physician in the medical domain is increasingly important. In medical domain diagnostic, classification and treatment are the main task for a physician. Today, Clinical Decision Support System (DSS) are developed to act multi-purposed and are combined with more than one Artificial Intelligence (AI) method and technique [16].

2.0 LITERATURE REVIEW

Soomro*et al.,* developed Knowledge Based Expert System for Symptomatic Automated Healthcare in 2011. The system is a knowledge based Expert System, having three chief modules-user interference engine and. Patient or users remotely interact with the system and find disease by giving, some symptoms to computer in this way; the systems make feasible diagnosis for patients, and also suggest particular treatment regarding the disease. The system is plague with the fact that there are many diseases under considerations; the accuracy of the system was not measured. A system with main focus on typhoid fever may provide better accuracy [17].

Vassal V. *et al.,* in 2008 solution trees were employed to diagnose salmonellosis, botulism, hepatite, leptospirosis, typhoid fever and dysentery. Complaints, objective data and laboratory checkup were used in the solution trees formation. The tree was constructed on the basis of experts' collective evaluation with the usage of paired comparison method. Algebraic methods of experts' information processing were used for the formation of an "effective" tree. The algorithm of consecutive analysis of variants was used for the determination of optional ways in solution trees that allows elaborating trees with thousands of peak. The task is becoming more complicated if not a "fixed" but a "dynamic" knowledge is needed "way of thinking"[18].

Sunday T. *et al* (2013). A Rule Based Expert System for Diagnosis of fever was developed. The research motivations include: the two most common forms of fever in Nigeria are malaria and typhoid fever.
The methods used in this work involved the development of expert system based on data collection and interaction with medical expert. The basis used for the generation of rules was not early stated[19].

Ital*et al.,*(2004) carried out a research on correlation studies on Widal Agglutination Reaction and Diagnosis of Typhoid Fever. In this research, 80 patients suspected of having typhoid fever infections were screened for the presence of Salmonella species using blood, urine and stool samples along with Widal agglutination tests. The result of statistical analysis revealed significant differences between the widal agglutination reaction and cultural diagnosis of clinical samples and strongly suggested that serological investigation alone may not be a reliable diagnosis for enteric (typhoid) fever infections. This shows the need to look beyond laboratory tests only in the diagnosis of typhoid fever [20].

Adhor2011 system combined the action oriented, IMCI (Integrated Management of Child Illness) and the Disease-Oriented HIS (Health Information System) approaches to diagnose malaria and typhoid fever. The system carried out its diagnosis based on signs and symptoms, but lay great emphasis on the fact that medicine is evidence based. Rapid prototyping, using a single expert system shell, was used to develop the system due to its simplicity and fast learning curve. Other approaches to diagnosis of typhoid fever as main subject matter were suggested[21].

Putu*et al* (2012), developed a Fuzzy knowledge Based System for the diagnosis of the tropical infectious diseases. The expert system designed in the study used fuzzy logic and certainty factors for the diagnosis. Malaria, Dengue fever, Typhoid fever and Chikunguaya were diagnosed with the system. The system carried out diagnosis without therapy[22].

Adehor (2008) developed an Intelligence Decision support System for the prompt diagnosis of malaria and typhoid in the malaria belt of Africa. The motivations for this work include: current diagnostic tools are affected by the harsh tropical weather, lack of qualified laboratory technicians, lack of regular supply of electricity to preserve diagnostic tools. A study was carried out which confirmed that both typhoid and malaria fever could be diagnosed based on signs and symptoms. A simplistic differential diagnostic model for the diagnosis of malaria, typhoid fever and unknown fever was formulated. The knowledge analysis of the system was carried out using mockler situation. The system was developed using rapid prototyping with simple expert system shell [23].

Prihatini (2013), Fuzzy expert system for tropical infectious disease by certainty factor was developed. The system used fuzzy logic to diagnose tropical diseases including typhoid fever. The system only carried out diagnosis, the therapy that could make it a perfect solution was neglected. Sunday T. (2013). A ruled Based Expert System for diagnosis of fever was developed for malaria, Typhoid, Dengue. The performance of the system was not measured to know its effectiveness. The chance of rules based system in handling cases not in knowledge base is slim. The severities of the symptoms were not put into consideration [24].

3.0 RESEARCH METHODOLOGY

Architectural Framework of the System

In this paper an intelligent ANN technique is adopted for the multiclass malaria and typhoid coinfection model. *Figure1* present system block diagrams for the development of data mining classification model: Artificial Neural Network (ANN).

Figure 1: Block diagram of Architectural System Framework

.

The data collected were classified and preprocessedthen subjected to training, testing and validation.

3.1 Description of the Data Set The records of patients were collected from General Hospital Mubi and Hong, in Adamawa State, Nigeria. These records of patients diagnosed withmalaria and typhoid fever were carefully selected and examined from medical practitioners.

The first seventy percent (70%) data sets, collected in September, 2016 were used for training set while another thirty percent (30%) collected in November, 2016 fifteen percent (15%) were used for testing and fifteen percent (15%) for validation in both Malaria and Typhoid.

A total of two hundred (200) data collected with malaria and one hundred and sixty (160) typhoid infection. The record consists of three (3) classes for typhoid infection and four (4) classesfor malaria infection according to the level of severity.Giemstastaining laboratory test was conducted on the blood samples.

3.2 Data Input

The symptoms that a patient observed are the present of bacterial (Salmonella Typhi and Paratyphi) in the body and the test results of the patient. It shows symbols and the range of their possible values called domains. Those symptoms for typhoid has three possible values which can be negative ($_0$), low ($_1$) and high ($_2$) while malaria has four (4) possible values which can be negative ($_0$) low ($_1$) mild ($_2$) high ($_3$).

4 DATA TRANSFORMATION

The data are classified based on the level of severity

Table 1: Data Transformation for Typhoid

Widal Count	Extent	Class
$x \leq \frac{1}{80}$	Not Significant	0
$\frac{1}{80} \leq x \leq \frac{1}{160}$	Low	1
$\frac{1}{160} \leq x \leq \frac{1}{360}$	High	2

From Table 1: A patient withwidalcount less or equal to $\frac{1}{80}$ signifies insignificant, a patient with widal count greater than $\frac{1}{180}$ or equal to $\frac{1}{160}$ is low class whilea patient with widal count more than or equal to $\frac{1}{360}$ is high.

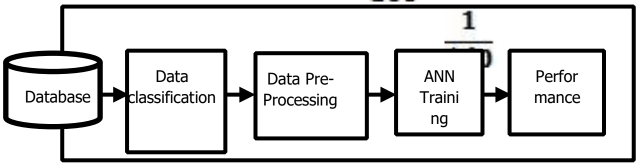


Table 2: Data Transformation for Malaria

MP	Extent	Class
-ve	Negative	0
+	Low	1
++	Mild	2
+++	High	3

From Table 2: A patient with negative (-ve) malaria parasite count is of class 0, a patient with a positive (+ve) malaria parasite count is of class 1, a patient with two positive (++ve) malaria parasite count is of class 2 while a patient with three positive and above (+++ve) malaria parasite count is of class 3.

5 Data Normalization and Description

Pre-processing with divide by maximum method was implemented on the data after removal of outliers. The input data of typhoid is the 200 samples with 8 attributes representing a 200 x 8 dataset also, the input data of malaria is the 160 samples with 7 attributes representing a 160 x 7 dataset.

Network Training

The network training comprised 70% of the proposed sample which is 140, 112 samples respectively. The processed data served as input to ANN. The network training also comprises of 2 layers,
8, 7 inputs, 10

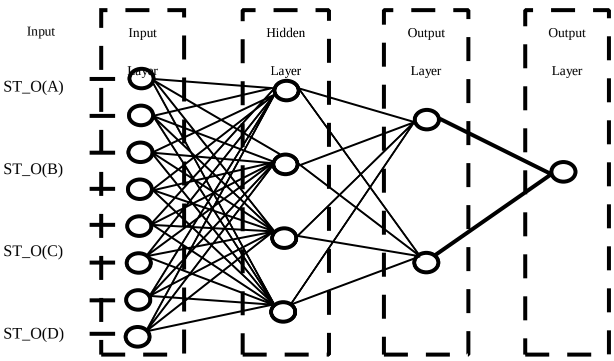
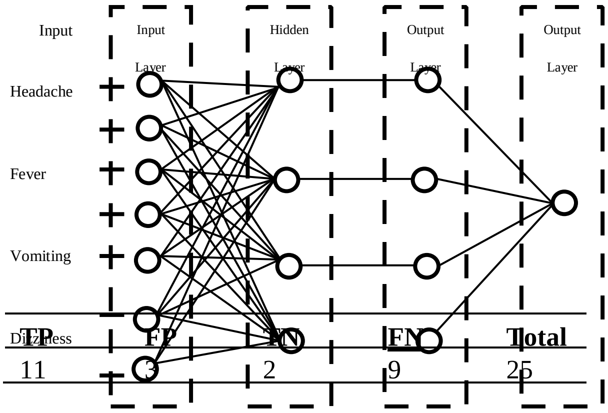


Figure. 2: Schematic Diagram of Typhoid Neural Network



TP	FP	TN	FN	Total
11	3	2	9	25

Figure. 3: Schematic Diagram of Malaria Neural Network

hidden layer and 10 hidden neurons feedback propagation ANN which was design for the training of typhoid and malaria data set using the processed data set. *Figure 2.*and*Figure 3* present Neural Network schematic description of Typhoid and Malaria infection.

Network Testing

The sample used for testing and validation is 15%, 15% of data respectively. In this paper, Sixty (60), and Fifty (50) of Typhoid and Malaria data were used for testing and validation. The network was tested and validated to ascertain its performance by obtaining the number of samples that were correctly classified when compared with target data.

4.0 RESULTS AND DISCUSSION

A true positive (TP), False Positive(FP), True Negative(TN), and False Negative(FN) performance metrics were used for the result analysis as shown in Table 3 and 5. Also the accuracy, sensitivity and specificity of the model were examined.

TP	FP	TN	FN	Total
14	0	0	50	64

Table 3: ANN matrix test result on typhoid cases

Table 3 shows the summary of the matrix results for typhoid samples data indicating TP, TN, FP and FN. The matrix obtained using the ANN was recorded accordingly, out of 14 data consider as positive by the physician, the ANN found that 14 were positive and 0 were negative. Similarly, out of the 50 data samples that the physician considered negative, the ANN found that 0 were negative and 50 were positive. Therefore, the ANN gave values of 22% accuracy, 12% specificity and 100% sensitivity as shown in Table 4. These results were obtained after several testing was performed on the ANN MATLAB tool.

Table 4: ANN Test Results

Performance Measurement	ANN Performance
Accuracy	22%
Specificity	12%
Sensitivity	100%

Table 5: ANN matrix test result on malaria processed data

The matrix obtained using the ANN is given in Table 5 accordingly, out of 14 data that the physician considered as positive, the ANN found that 11 were positive and 3 were negative. Moreover, out of the 11 data that the physician considered negative, the ANN found that 2 were negative and 9 were positive. Therefore, the ANN gave values of 62% accuracy, 57.89% specificity and 100% sensitivity as shown in table 4.8. These results were obtained after several testing was performed on the ANN MATLAB tool
(Class 1 was choice the best).

Table 6: ANN Test Results

Performance Measurement	ANN Performance
Accuracy	62%
Specificity	57.89%
Sensitivity	100%

ANN Classification Test Results

Table 4 and Table 6 show some of ANN Test Result on the 200 and 160 of typhoid and malaria normalized data. Table 3 and 5 shows the summary of the ANN matrix test results for typhoid and malaria samples data indicating the TP, TN, FP and FN.

5.0 CONCLUSION AND RECOMMENDATION

This study developed data mining techniques for classification of typhoid and malaria disease using ANN in MATLAB environment. The first objectives are to classify data sets, and transform it into preprocessed data. The network was created and the data was train with the network. The performance of the system was measured on both the training set and testing data. A MATLAB/Excel was used for the machine to learn. The accuracy, specificity, and sensitivity matrix are very important to better evaluate the performance of a technique which enables the researcher to achieve its objectives. This study used a data mining classification technique for the classification of typhoid and malaria disease datasets. It is recommended for further studies to use large volume and different dataset for the classification of all the diseases. Contributions and Suggestions are welcome at this stage of the research.

REFERENCES

[1] Shiadeh, M. N., Niyyati, M., Fallahi, S., &Rostami, A. (2016). Human parasitic protozoan infection to infertility: a systematic review. *Parasitology research*, *115*(2), 469-477.

[2] Spira, A. M. (2003). Assessment of travellers who return home ill. *The Lancet*, *361*(9367), 1459-1469.

[3] Owoseni, A. T., & Ogundahunsi, I. O. (2016). Mobile-Based Fuzzy Expert System for Diagnosing Malaria (MFES). *International Journal of Information Engineering and Electronic Business*, *8*(2), 14.

[4] Tunmibi, S., Adeniji, O., Aregbesola, A., &Dasylva, A. (2013). A rule based expert system for diagnosis of fever. *International Journal of Advanced Research*, *1*(7).

[5] Winstanley, P. A. (2000). Chemotherapy for falciparum malaria: the armoury, the problems and the prospects. *Parasitology Today*, *16*(4), 146-153.

[6] Wilson, M. E., Kantele, A., &Jokiranta, T. S. (2011). Review of cases with the emerging fifth human malaria parasite, Plasmodium knowlesi. *Clinical infectious diseases*, *52*(11), 1356-1362.

[7] Bhan, M. K., Bahl, R., &Bhatnagar, S. (2005). Typhoid and paratyphoid fever. *The Lancet*, *366*(9487), 749-762.

[8] Guzman, C. A., Borsutzky, S., Griot-Wenk, M., Metcalfe, I. C., Pearman, J., Collioud, A. & Dietrich, G. (2006). Vaccines against typhoid fever. *Vaccine*, *24*(18), 3804-3811.

[9] Samuel, O. W., Omisore, M. O., &Ojokoh, B. A. (2013). A web based decision support system driven by fuzzy logic for the diagnosis of typhoid fever. *Expert Systems with Applications*, *40*(10), 4164-4171.

[10] Gupta, V., &Garg, R. (2014). Typhoid Fever and Water. In *Water and Health* (pp. 93-106). Springer India.

[11] Nuccio, S. P., Wangdi, T., Winter, S. E., &Bäumler, A. J. (2013). Typhoid. In *The Prokaryotes* (pp. 375-399). Springer Berlin Heidelberg.
[12] Oguntimilehin, A., Abiola, O. B., &Olatunji, K. A. (2015). Computer Aided Diagnostic Systems for Managing Typhoid Fever: A Review of Diagnosis Techniques. *International Journal of Computer Applications*, *126*(6).
[13] Ansumana, R., Jacobsen, K. H., Gbakima, A. A., Hodges, M. H., Lamin, J. M., Leski, T. A., ...&Stenger, D. A. (2013). Presumptive selfdiagnosis of malaria and other febrile illnesses in Sierra Leone. *Pan African Medical Journal*, *15*(1).
[14] Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in medicine*, *23*(1), 89-109.
[15] Jackson, J. (2002). Data mining; a conceptual overview. *Communications of the Association for Information Systems*, *8*(1), 19.
[16] Ahmed, M. U. (2011). *A Multimodal Approach for Clinical Diagnosis and Treatment* (Doctoral dissertation, Mälardalen University).
[17] Soomro, A. A., Memon, N. A., &Memon, M. S. (2011). Knowledge based expert system for symptomatic automated healthcare. *Sindh University Research Journal-SURJ (Science Series)*, *43*(1 (a)).
[18] Vassal Vassileu et al (2008), "Diagnostic Systems in Medicine As Personal Intellectual Tooling", International Journal "Information Technologies and Knowledge", Vol.2/2008 .Pg 211-217.
[19] Sunday Tunmibi, OriyomiAdeniji, Ayooluwa Aregbesola and Ayodeji Dasylva (2013), " A Rule Based Expert System for Diagnosis of Fever", International Journal of Advanced Research, vol1, issue 7, pg 343-348. www.journalijar.com. 12/03/14.
[20] Alfred *et al* (2004) carried out a research on correlation studies on Widal Agglutination Reaction and Diagnosis of Typhoid Fever
[21] Iketut G.P and Putu M.P., " Fuzzy Expert System for Tropical Infectious Diseases by Certainty Factor", TELKOMNIKA, Vol.10, No4, pg 825-836.
[22] Adehor, A. B., & Burrell, P. R. (2008, September). An Intelligent Decision Support System for the Prompt Diagnosis of Malaria and Typhoid Fever in the Malaria Belt of Africa. In *IFIP International Conference on Artificial Intelligence in Theory and Practice* (pp. 287296). Springer, Boston, MA.
[23] Prihatini, P. M. (2012). Fuzzy Expert System for Tropical Infectious Disease Diagnosis by Certainty Factor. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *10*(4), 825-836.

AUTHORS' PROFILES

Abisoye Opeyemi A. was born in Ogbomoso, Oyo State, Nigeria. She attended University of Ilorin, Ilorin, Nigeria where she obtained her BSc, Msc, degree in Computer Science,. She is currently a PhD. Student of the same institution. She is major in Computational Intelligence, Machine Learning, Data Mining, and Soft Computing. She serves as a Lecturer I, in the Department of Computer Science, SICT, Federal University of Technology, Minna, Niger State, Nigeria from May 23rd 2007 till Date.

Professional Membership: She is a registered member of Computer Professionals [Registration Council of Nigeria]-**MCPN (30th June,).**

Douglas Ibrahim was born in Gombi, Adamawa State, Nigeria. He attended Federal Polytechnic Mubi, Adamawa Nigeria where he obtained hisHND in Computer Science; He is currently a PGD student of the Federal University of Technology Minna. He serves as anInstructor II, in the Department of General Studies, College of Agriculture, Ganye, Adamawa State, Nigeria.

Abisoye Blessing O. was born in Ogbomoso, Oyo State, Nigeria. He attended University of Ilorin, Ilorin, Nigeria where he obtained his B.Eng., MEng., degree in Electrical and Electronic Engineering,. He is major in Power and Control Engineering.
He serves as an Assistant lecturer in the department of

Computer Engineering, Federal University of Technology, Minna, Niger State, Nigeria from May 23rd 2012 Till Date.

Professional Membership: He is a Registered member of Nigeria Society of Engineers (NSE) and Council for the Regulation of Engineering in Nigeria (COREN)

Elisha Richard was born in Hong, Adamawa State, Nigeria. He attended Federal Polytechnic Mubi, Adamawa, Nigeria where he obtained his HND in Micro Biology; He is currently a PGD student of the Adamawa State University Mubi.He serves as aTechnologist I, in the Department of Science and Laboratory Technology, Federal