

Case-Based Reasoning Framework for Malaria Diagnosis

Eshetie Gizachew Addisu

School of Computing, Ethiopian Institute of Technology-Mekelle, Mekelle University, Mekelle, Ethiopia
E-mail: eshexas11@gmail.com

Abiot Sinamo Boltena

School of Computing, Ethiopian Institute of Technology-Mekelle, Mekelle University, Mekelle, Ethiopia
E-mail: abioticsinamo35@gmail.com

Samson Yohannes Amare

School of Computing, Ethiopian Institute of Technology-Mekelle, Mekelle University, Mekelle, Ethiopia
E-mail: samsonya@gmail.com

Received: 14 March 2020; Accepted: 06 September 2020; Published: 08 December 2020

Abstract: Malaria is life threatening disease in Ethiopia specifically in Tigray region. Having common symptoms with other diseases makes it complex and challenging to diagnose effectively. In this paper case based reasoning framework for malaria diagnosis has been designed to diminish the challenges faced by inexperienced practitioners during malaria diagnosis and to solve the problem on shortage of health professionals. The required knowledge for this study was collected through interview and document analysis from domain experts, malaria patient history cards and other related relevant documents. In the case acquisition process the manual format of cases makes the process too challenging. Decision tree is used to model the acquired knowledge. The case structure was then constructed using the selected most determinant attributes. Machine learning approach is applied to select the most relevant features. Feature-vector case representation technique is applied to represent the collected malaria cases. Jcolibri programming tool integrated with Eclipse and Nearest Neighbor retrieval algorithm are used to design the framework. To the end based on the results we can say that the machine learning approach can be used to select most relevant attributes in diseases having several common symptoms and designing case-based diagnosis frameworks could overcome the main problems observed in health centers of Tigray. As an artifact the framework is evaluated by statistical analysis, comparative evaluation, user evaluation and other evaluation techniques. Averagely 79 % precision, 89 % recall, 91.4% accuracy and 78.8% domain expert's evaluation was the results scored.

Indexing Terms: CBR, AI, Case Representation, Framework, knowledge base systems, Attribute selection, Diagnosis and Treatment Guidelines.

1. Introduction

Malaria is one of the major worldwide leading public health problems [1]. In Ethiopia, more than half of the total population is living in malaria endemic area which makes them victim of malaria [2]. It is caused by a protozoan of the genus plasmodium. Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae and Plasmodium ovale are well known species of malaria which infect and cause disease in human. Malaria poses a tremendous public health problem and significantly affects the poor society, who suffers socio economic problems [1]. This is happening due to shortage of professionals and scarcity of laboratory equipment especially in less developed countries. Thus, it is difficult to have satisfactory diagnosis and treatment of malaria with these kinds of situations. This research focuses on automating and modernizing malaria diagnosis by applying artificially intelligent (case-based reasoning) techniques in Ethiopia, particularly in Tigray region. The main objectives in this study are to design a Case-Based Reasoning Framework for malaria diagnosis by selecting the most determinant features to support health professionals and improve the service provided and to evaluate it.

KBS is one of AI technologies and it can be defined as a computer program that represents and reasons with knowledge of some specialist subject and documents with a view of solving problems or giving an advice. It is a type of application programs that plays a great role in health care institutions. So far, several intelligent systems are developed by different researchers to facilitate and support diseases diagnosis and as decision support systems for health professionals. At the early stage of medical intelligent systems, when the implementation of MYCIN [4] most of the

developed systems, especially knowledge base diagnosis systems had attempted to apply rule-based technique [5]. Later, applying the rule based system for a broad and complex medical domain faces several problems and limitations through time [6]. A rule-based system are good enough if the problem domain is well understood, constant over time and the domain theory is strong[6].

Even though, there are well experienced health professionals in malaria diagnosis, it is difficult to articulate and express their knowledge explicitly in case of rule-based systems [6]. Large differences of individuals (experts) in the same problem domain, the influence of environmental background of the patient and the absence of general rules make it difficult to diagnose disease (malaria) and predict the risk of malaria-related health problems [6]. On the other hand, Case-Based Reasoning (CBR) approach works well in such domains where the domain knowledge is not clear enough. Many systems (knowledge base systems and expert systems) were developed employing different techniques, in the attempt to improve the quick and safe establishment of a diagnosis including MYCIN. These systems are mainly characterized by the manipulation of symptoms and other measurements by applying rule-based technique. To enhance the performance of the rule based systems, researchers recommended several future research works such as to apply other reasoning techniques (Case-Based Reasoning, Neural Network, Hybrid and Ontology). CBR can optimize the diagnosis systems and helps the KBS to learn from the previously solved diagnosis documents (cases).

Tigray is one of the regions located in the northern part of Ethiopia where malaria is highly prevalent. Plasmodium falciparum and plasmodium vivax are the most common species in this region [7]. Even though, malaria is curable by effective early diagnosis and treatment, the ratio of malaria patients and available health care centers in Tigray is unbalanced [8]. Malaria could lead to death, if it is not treated at its initial stage. In diagnosing malaria, medical decision making is also another challenge encountered for inexperienced professionals. Because the symptoms observed in different patients is varying and the degree of determinations of each symptoms towards malaria diagnosis is another factor affecting the decision making task. So, determining the important symptoms and parameters for malaria diagnosis and designing knowledge base framework (case based reasoning frameworks for malaria diagnosis) to the available health care centers and hospitals of Tigray will alleviate the decision making problem, lack of experienced professionals and improves the quality of service provided.

There are medical systems developed so far in Ethiopia and even in Tigray to support malaria diagnosis and treatment by Chala[2] and Aster[3]. Chala[2] attempt to develop learning KBS for diagnosis and treatment of Malaria and Aster[3] also try to develop local expert system for diagnosis and treatment of malaria. Even though, Chalas'[2] work is better in terms of its learning capabilities, these studies are mainly limited to apply case based reasoning technique and any approach to select the most determinant factors considered in malaria diagnosis. In addition to that both Chala[2] and Aster[3] simply acquire the domain knowledge and design the system by using the rigid rule based reasoning technique. Therefore, based on the gaps identified on these studies; the researchers of this study hopes to identify the most determinant features in malaria diagnosis and to design CBR framework for malaria diagnosis for health centers in Tigray to improve the malaria diagnosis service provided. To the end, the following research questions are assessed and answered.

1. What are the most determinant symptoms and parameters in malaria diagnosis?
2. What is the optimal architectural representation of the case based reasoner?
3. How to transform the architecture to a usable artifact?
4. How to design the components to provide the case based reasoning framework?

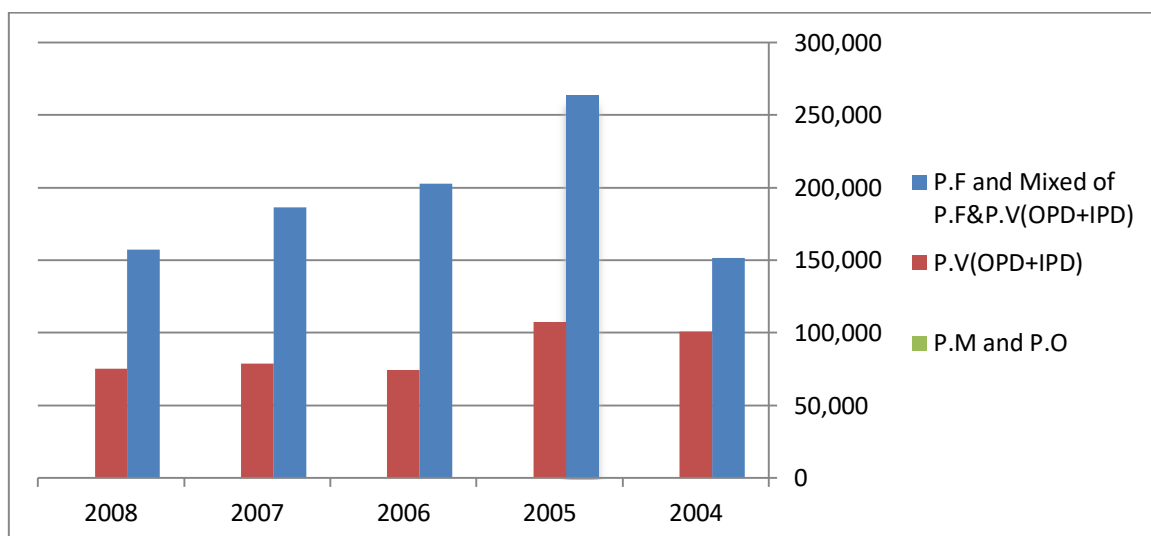


Fig.1. Malaria species spread in Tigray region. Source (-7-)

Even though there are four well known malaria parasite types worldwide, in Tigray only the two parasites are common as depicted in the above figure 1.

2. Methodology

2.1. Literature Review

Several researches have been conducted to investigate the applicability of knowledge based (CBR) systems in different domains all over the world including in our country Ethiopia. Locally case-based reasoning technique is applied in health domain by different researchers such as; Alemu [9] investigated the potential of case based reasoning in solving complex side effects of HIV/AIDS cases for person living with HIV/AIDS who have begun antiretroviral therapy. He used JCOLIBRI version 1.1 in designing the prototype. He uses ten (10) test cases and the system registers 72% and 63% of recall and precision respectively [9].

Henok[8] developed a prototype knowledge based system using CBR technique for hypertension management. He used 45 hypertension patient cases for building and testing the prototype by using seven (7) test cases from the case base. Henok[10] registered 86.1% recall and 60% average precision [10]. Getachew [5] had also done research study on application of case-based reasoning for anxiety disorder diagnosis. The testing of the prototype is performed from two sides. The first one is testing in terms of precision and recall and registered 71% and 82% respectively [9]. The second one is the performance of the system is evaluated by the potential users' of the system and achieved 83.2% performance.

Moreover there have been many CBR systems developed in medical domain to support different disease diagnosis abroad.

CASEY: is a knowledge based system combines case-based and rule-based techniques to support heart failure diagnosis using three steps; search for similar cases, determination of differences and their evidences, and a transfer of the diagnosis of similar cases [11]. Even though attempting to use more general adaptation was the interesting aspect of CASEY, all adaptation problems could not be solved.

ALEXIA: is case-based reasoning which used a deep pathophysiological model of hypertension to assess the similarity between cases favoring the most important attributes in the model [12].

MNAOMIA: is a case-based reasoning system that can adapt to the medical task at hand, namely diagnosis, treatment planning, and clinical research assistance [13]

CARE-PARTNER: is a case-based reasoning system proposed a cooperation framework between cases and clinical practice guidelines in the domain of stem-cell post-transplant care [14].

Recently local researchers tried to apply the CBR technique to different domain areas to overcome the limitations of rule based systems such as; Biazen [15] investigated on the application of case based recommender system in field of study selection in the case of higher education in Ethiopia with main objective of assisting students in their field of study selection. He used one hundred five (105) old cases which are collected from history of successful students as case base. The prototype of this system is tested for its performance and user acceptance evaluation and it registers an average performance of 72% by domain experts and 80.2% by students [15]. Yibeltal [16] has also investigated case based reasoning recommender system in the investment domain with the objective of recommending and assisting the foreign and domestic investors in selection of investment sectors and investment activity. Finally the user acceptance of this case based recommender system had been tested to evaluate its performance and, it achieves 82% and 84% by the domain experts and investors respectively [16].

Generally, almost all researchers who investigated case based reasoning in different domain area faces similar challenges, since the retrieval algorithms and case representation used are similar. In addition to this most of the local authors recommend that identifying determinant parameters (attributes) of cases and increasing the number of cases will improve the performance of the CBR systems. Therefore, the aim of this study is to explore the applicability of case based reasoning frameworks in health domain to improve the decision making problem and enhance the process of malaria diagnosis in Tigray by increasing the number of cases and selecting most relevant attributes to check the effect on CBR framework for malaria diagnosis in all CBR performance evaluation aspects.

2.2. Data collection

The data collection process of this study involves the document analysis from different sources and interview with domain experts in Kahasay Abera and Ayder Comprehensive Specialized Hospitals. Mainly to acquire knowledge of the domain experts an interview is conducted as a means of detail discussion among the researcher and experts. There are several interview techniques which can be used differently to acquire knowledge and collect facts while semi-structured type of interview is selected and applied for this research. The main reason for selecting semi-structured interview is it allows the researcher to acquire detail domain knowledge by raising another interview questions when he/she needs more clarification in between. Since this style of interview incorporates both structured and unstructured interview techniques it allows subsequent questions. Thus an in-depth semi-structured interview is conducted with the domain experts to know the main challenges faced in diagnosing malaria, to identify the main symptoms used to detect malaria parasites, to understand about types and spread of malaria parasite types and to understand other related facts on malaria diagnosis process.

Moreover, document analysis is also conducted to understand about the domain knowledge in addition to collecting the past malaria patient history cases from the patient history record document. Again after malaria patient cases are collected the domain experts are consulted for the purpose of more clarification and verification of the collected cases. In the time of collecting cases the role of card room experts and Health Management Information Systems (HMIS) officers in searching and recording the patient history to the computer system was irreplaceable. Here the first task was identifying malaria patient registration number from the HMIS office, then after manual searching of the patient history from the shelf and recording to the computer system was the next task.

2.3. Sampling Technique

Though there are several sampling techniques suitable for different research methodologies and domain areas purposive sampling technique is selected for this study. Thus in the knowledge acquisition process domain experts are selected to be interviewed purposely based on different criteria such as experience in the domain, academic rank and others. In addition to this the selection of past malaria cases to build the case base and selection of test cases in evaluation phase is also conducted purposely based on their completeness, time, relevancy and other criteria.

2.4. Knowledge Modeling

Once the knowledge is acquired from different sources, knowledge modeling is applied to represent the real process followed during malaria diagnosis in understandable way. In this study decision tree modeling technique is used to model the acquired knowledge.

2.5. Case Representation

Knowledge representation formalism (case representation) in case based reasoning is used to represent acquired knowledge contained in the collected cases for the purpose of reasoning. There are several varieties of case representation formalisms such as, feature-value, structured, textual and others, but the researchers of this study selects feature-value case representation. It represents cases in suitable way to be easily understood by the programming tools selected. So, the collected cases are going to be represented as feature-value representation with the selected attributes for the efficient retrieval process in the programming tool.

2.6. Implementation Tools

jCOLIBRI2 integrated with eclipse Integrated Development Environment (IDE) is used for the implementation of this study.

3. Conceptual Modeling

After acquiring the needed knowledge from domain experts, malaria patient cases and other relevant related documents, conceptually modeling the acquired domain knowledge was the proceeding task. Organizing and structuring knowledge gathered during the knowledge acquisition are the main steps involved in knowledge modeling. This activity gives an implementation independent specification of the knowledge to be represented in the designed knowledge based framework. Knowledge modeling is critical stage of knowledge acquisition and it can be expressed as the concept of representing information and the logic for purpose of capturing, sharing and processing of knowledge to simulate intelligence [18]. It helps to model the main concepts that tell the basic activities and decisions made to solve cases in the problem domain. Conceptual modeling enables to understand the problem domain and to prepare the knowledge representation phase. It can be modeled by applying any conceptual modeling techniques, but for this study decision tree structure is used to model how the overall activities are performed during malaria diagnosis.

Decision tree is most extensively applied methods of conceptual modeling and commonly it plays a great role in a knowledge modeling process. In medical domain decision tree is well known tool and represented by graph. After all, the conceptual model for the case based reasoning framework is modeled to make the acquired knowledge functional in knowledge representation by using the decision tree. The design of case based reasoning framework for malaria diagnosis follows the same procedures as presented in the decision tree. During the interview of knowledge acquisition process, domain experts had explained that there are two main factors considered for diagnosing new patient who came for malaria diagnosis. These are concepts of symptoms and concepts of laboratory test.

As we try to model in the below figure 2, the actual malaria diagnosis process in Tigray health institutions is conducted as presented in figure. According to the domain experts explanation in the interview during malaria diagnosis knowledge acquisition process ,first of all health professionals would have a question and answer with the patients or their relatives on what the patient feels and experiences in the past hours or days. If the patient is suspected to be malaria attacked then blood test will be conducted by taking the blood films. The blood laboratory test could be done by using either the microscope or the Rapid Diagnostic Tests (RDT). Based on the laboratory result the specific malaria parasite would be detected and treated accordingly if the patient is being infected by falciparum or vivax malaria parasites. But, if the patient had symptoms of malaria in the question and answer with health professionals but neither of

malaria parasite is diagnosed, the patient will sent for further diagnosis in case if he/she have other malaria parasites. But as depicted in the figure 1, there is no other malaria parasites found in Tigray than falciparum and vivax parasites.

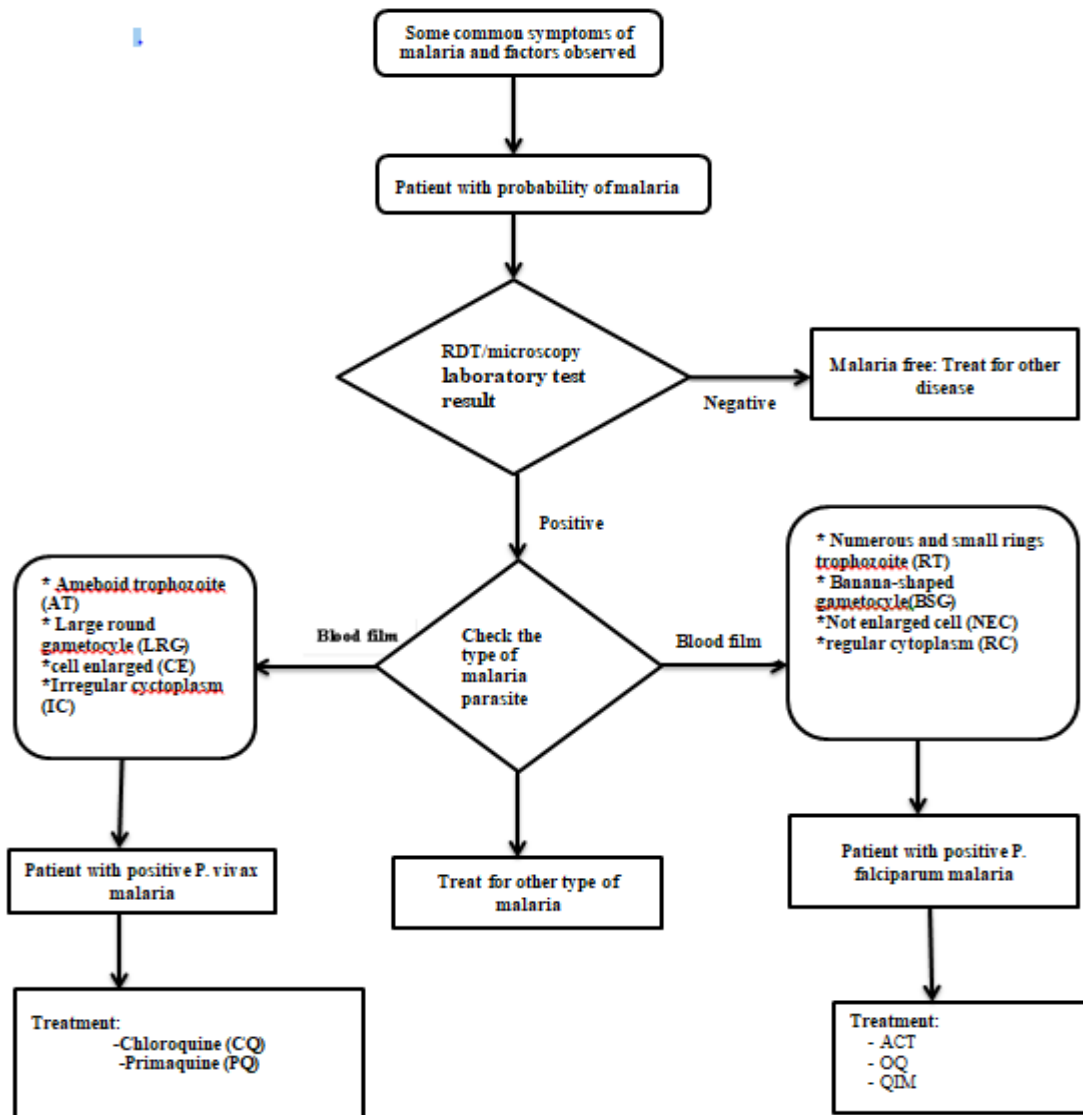


Fig.2. Decision tree for malaria diagnosis

4. Malaria Diagnosis Case Structure

Cases in case based reasoning frameworks includes two main parts which are known as problem description part and solution part. As case based reasoning frame work, cases of this study also have both problem description and solution part. The problem description part sometimes also called situation.

Problem description/Situation: this is a part of the case structure which contains attributes of malaria case.

Solution: this part of case structure justifies the type of malaria parasite detected and the way malaria diagnosis is treated based on the result of problem description part.

The identification of attributes which comprises both problem description and solution parts is performed by the researcher with the help of domain experts. All attributes are identified and selected from the patient history card record but attributes to design the case structure are finally selected by applying information gain machine learning approach for attribute selection purpose. Even though there are a lot of attributes in the record, only attributes which have high relevancy based on information gain values of attributes in the malaria diagnosis are selected. The cases that build the case base are collected from malaria patients' card history in the case acquisition phase are shaped to the constructed case structure. This stage was too challenging, because all patients of any disease including malaria card history is kept in a single room so it is a tedious task to find malaria patient card from these millions of patient card histories. Not only

searching was the challenge, but even after getting malaria patient card history there are also other challenges. For instance some patient card history lacks the treatment and other necessary attributes such as the blood laboratory result gained due to the carelessness of the physician's. Another difficult challenge here was converting the hard copy contents of the case in to soft copy based on the constructed case structure, because the patient card history are available in paper format which is impossible to build the computer based framework to diagnose disease with such types of files.

Table 1. Case Structure of Malaria Diagnosis Cases

Attribute Name	Parameter
Age	Problem Description
Sex	Problem Description
Pregnancy	Problem Description
Sweating	Problem Description
Fever	Problem Description
Headache	Problem Description
Anemia	Problem Description
Weakness	Problem Description
Vomiting	Problem Description
Myalgia	Problem Description
Chills	Problem Description
Shiver	Problem Description
Loss of appetite	Problem Description
Travel history within 30 days	Problem Description
From Malaria endemic Area	Problem Description
Lab Result	Problem Description
Assessment	Solution
Treatment	Solution
Explanation	Solution

5. Architecture of Cbr Framework for Malaria Diagnosis (Cbrfmd)

This section mainly involves the task of designing CBR framework for malaria diagnosis by using an appropriate programming tool and algorithm. For this study jCOLIBRI integrated with eclipse and Nearest Neighbors are the tools and retrieval algorithm used to design the case based reasoning frame work respectively. Eclipse is a java editor used with jCOLIBRI for writing and executing the java codes. The main reason to select jCOLIBRI is that it uses nearest neighbor retrieval algorithm for the retrieval of cases from case base which has similarity with the new case (query) and it is also suitable when there are attributes with numeric (continuous) values [19]. The figure 3 below shows the architecture of the CBRFMD which precisely depicts the way the framework is designed and handles malaria diagnosis in health care centers. As presented in the architecture the new patient cases fed to the user interface as a query, then the framework searches and retrieve match from old solved cases (from the case base) by using similarity measurement. The relevant retrieved cases from the case base are ranked based on their global similarity when the relevant matched cases to the new case are more than one. Finally solution for the new case is proposed either by adapting (when the retrieved case and new cases match partially) or derive (when the retrieved case match with the new case perfectly) from the old case. At the end the proposed solution is revised for any inconvenience and then the revised solution is retained in the case base for future problem solving. So the case base updates incrementally when the framework learns from new case used by the users.

However there might be a situation of new case could not match with any case from the case base which means there is no similarity between the existing case and new case (query). This happens when there are no previous stored cases in the case base having any similarity with the new case (query) in all attribute values. Therefore if there is no similarity between the source and new case, the new case of the patient will be diagnosed out of the framework by the domain experts and then the case with its solution should be revised and stored in the case base. Finally, the revised solution or solved cases is retained in the case base for future problem solving as experience for similar cases.

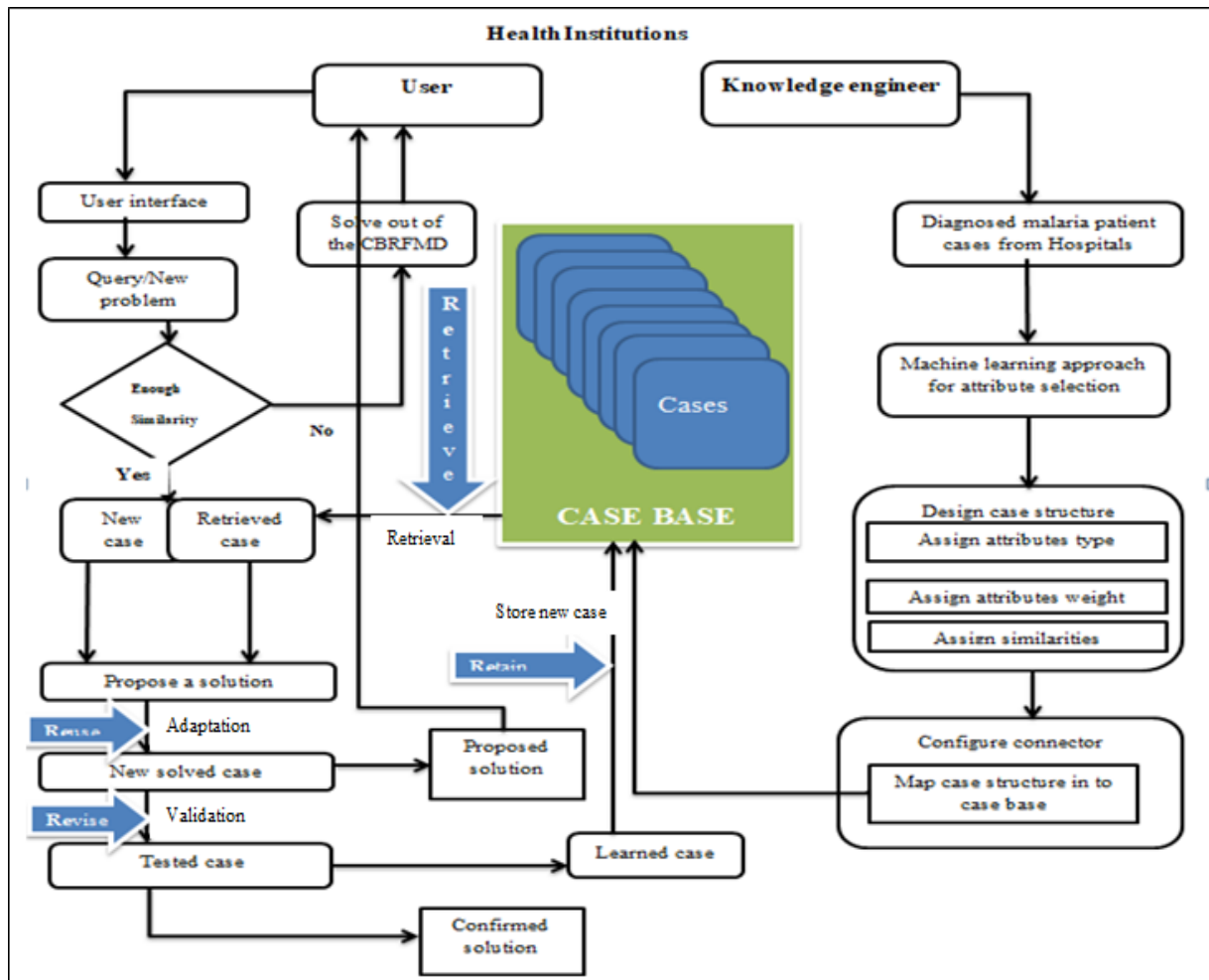


Fig.3. CBRFMD architecture

5.1. CBR Framework for Malaria Diagnosis (CBRFMD)

As a CBR approach the design of Case Based Reasoning Framework for malaria diagnosis was complex task which involves several steps, like acquiring solved cases and general knowledge, selecting best attributes, representing cases by applying appropriate case representation method, defining similarity measure, implementing retrieval functionality, implementing attractive user interfaces and soon[20].

A. Building the Case Base

Feature-value representation method is the selected technique to represent the cases to build the case base of malaria cases. So, collected cases for this study are stored as plaintext file in a feature-value representation format. An easy way of case representation and good integration with nearest neighbor retrieval algorithm are the main reasons to select feature-value case representation for this research [21]. Therefore the case representation is given as a plaintext feature-value representation consisting of N columns representing case attributes (A1, A2, A3... AN) and each M rows representing individual cases C ({C1, C2, C3, ..., CM}). Each attribute has a sequence of possible k values associated to each column attribute $A = \{V1, V2, V3, \dots, Vk\}$.

B. Attribute Selection

The attributes collected from malaria complaint history records are the body of the case based reasoning framework to perform malaria diagnosis. But all the attributes may not equally relevant for the decision making related to malaria diagnosis. The inclusion of less or not relevant attributes in automated activity operations such as case based reasoning framework for malaria diagnosis may degrade the accuracy of the framework. Due to this after past malaria cases for this research have been collected machine learning approach called information gain algorithm is applied to select relevant attributes for malaria diagnosis. Attribute selection helps to overcome over fitting and curse of dimensionality problems in addition to time and space complexity improvements. Information gain is a statistical measurement in bits to measure how much a single attribute on the instance classifies or separates the instance itself to the target classification. For this study information gain measure for each attribute is calculated and attributes with high

information gain are selected to represent instances of malaria patient cases. To do this a measure called entropy is used to calculate the information gain of each attribute. The entropy for a set S is calculated as:

$$\text{Entropy}(S) = \sum_{i=1}^n -p_i \log_2 p_i \quad (1)$$

Where n is number of classes, p_i is the probability of S belongs to class i. After entropy is calculated it is easy to find the information gain of any arbitrary attribute A as follow:

$$\text{Gain}(A) = \text{Entropy}(S) - \sum_{k=1}^m |S_k|/|S| (\text{Entropy}(S_k)) \quad (2)$$

Where S_k is the subset of S.

To design the case based reasoning framework for malaria diagnosis, seventy (70) old solved cases with three classes are used to build the case base and 36, 23 and 11 from the total tuples belongs to P.Viavax, P.Falciparum and Mixed of the two classes respectively. So to calculate the information gain of each attribute for the sake of attribute selection for this study by using the above formula it can be expressed with values: $S = 70 = (36, 23, 11)$, but, before calculating the information gain of each attribute it is mandatory to find entropy (S).

$$\begin{aligned} \text{So, entropy}(S) &= \sum_{i=1}^n -p_i \log_2 p_i \\ &= -36/70 \log_2 (36/70) - 23/70 \log_2 (23/70) - 11/70 \log_2 (11/70) \\ &= 0.494 + 0.528 + 0.420 \\ &= \mathbf{1.442} \end{aligned} \quad (3)$$

Here the formula to find the information gain of each attribute is applied as follow. For instance

Gain (Attribute sex)?

From the 70 malaria compliant records attribute sex is found with 25 'male' values and 45 'female' values with respect to the following classes.

$$\left\{ \begin{array}{l} C(P.V, \text{male}) = 10 \\ C(P.F, \text{male}) = 10 \\ C(P.V \& P.F, \text{male}) = 5 \end{array} \right\} \quad \text{and} \quad \left\{ \begin{array}{l} C(P.V, \text{female}) = 24 \\ C(P.F, \text{female}) = 14 \\ C(P.V \& P.F, \text{female}) = 7 \end{array} \right\}$$

Therefore, **Info Gain (Sex)?**

$$\begin{aligned} \text{Info Gain (Sex)} &= \text{Gain}(A) = \text{Entropy}(S) - \sum_{k=1}^m |S_k|/|S| (\text{Entropy}(S_k)) \\ &= 1.442 - 25/70 [-10/25 \log_2 (10/25) - 10/25 \log_2 (10/25) - 5/25 \log_2 (5/25)] + \\ &\quad 45/70 [-24/45 \log_2 (24/45) - 14/45 \log_2 (14/45) - 7/45 \log_2 (7/45)] \\ \text{Gain (Sex)} &= \mathbf{0.022} \end{aligned} \quad (4)$$

In this way the information gain of all attributes is computed and the result is depicted in the table 2 below.

Table 2. Information Gain Value Computed for All Attributes.

Attribute name	Gain
Sex	0.022
Age	0.102
Pregnancy	0.053
Sweating	0.057
Fever	0.089
Headache	0.02
Anemia	0.009
Weakness	0.113
Vomiting	0.248
Myalgia	0.017
Chills	0.331
Shiver	0.044
Appetite loss	0.006
Travel History	0.210
From Malaria Endemic Area	0.210
Lab Result	1.442

As shown in the table 2 almost all attributes have much better information gain value for the classification of malaria cases to their target class except two attributes which are with information gain value of 0.009(Anemia) and 0.006(Loss of appetite). Thus, the information gain value of these two attributes is almost negligible value relatively to the others attributes value. Standing from this numerical measurement attributes 'Anemia' and 'loss of appetite' are rejected from the design of case structure and case base for this study due to their less relevancy for malaria diagnosis framework design.

C. Case Representation

Knowledge representation formalism (case representation) in case based reasoning is used to represent experience knowledge contained in the cases for the purpose of reasoning. There are several variety of case representation formalisms such as, feature-value, structured, textual and others, but the researcher of this study selects feature-value case representation because feature value case representation represents cases as attribute-value pairs, similar to the propositional representations used in Machine Learning (ML), that support k-nearest neighbor matching and instance-based learning [11]. Feature-value case representation also represents cases in suitable way to be easily understood by the programing tool (jCOLIBRI).

Defining the features available in the case and measuring the similarity between existing cases and new case (query) is easily performed by defining case structure. As a case based reasoning framework this study is expected to retrieve similar cases from the case base for the query to support the decision making of professionals in health institutions. So, the collected cases are going to be represented as feature-value representation with the selected attributes for the efficient retrieval process in the jCOLIBRI programming tool and this is done through case indexing process. Indexing is defined as assigning indexes to cases for the purpose of retrieval and comparison of a new query to the case base [22].

D. Defining/Designing Case Structure

Once the feature-value representation is selected the whole collected cases are saved in plain text file format with respect to the selected attributes. The selected attributes of cases which are considered to have significant impact at the time of Malaria diagnosis are Sex, Age, Pregnancy, Sweating, Fever, Headache, Weakness, Vomiting, Myalgia, Chills, Shiver, Travel History, From Malaria Endemic Area and Lab result. String and integer data types are the main data types used to define different attributes of cases in this study. The significance of selected attributes in the malaria diagnosis is varying and their significance is assigned with the help of domain experts ranging from 0.1 to 1.0.

For the CBRFMD the description of case attributes by their name, data type, weight, local and global similarity is give as follow in the table 3 below.

Table 3. Descriptions of Selected Attributes and Solution Attributes

List of Attributes			
Attribute Name	Data Type	Weight	Local Similarity
Sex	String	0.5	Equal
Age	Integer	0.4	Interval
Pregnancy	String	1.0	Equal
Sweating	String	0.9	Equal
Fever	String	1.0	Equal
Headache	String	0.9	Equal
Weakness	String	0.9	Equal
Vomiting	String	0.9	Equal
Myalgia	String	0.8	Equal
Chills	String	1.0	Equal
Shiver	String	1.0	Equal
Travel History	String	1.0	Equal
From Malaria Endemic Area	String	0.8	Equal
Lab result	String	1.0	Equal
Solution Attributes			
Assessment	String	1.0	Equal
Treatment	String	1.0	MaxString
Explanation	String	1.0	MaxString

5.2. Similarity Measurement

The nearest neighbor similarity measurement in case based reasoning determines how much the new case and old cases are similar and it can be calculated by local and global similarity functions.

A. Local Similarity Measurement

Local similarity helps to measure the similarity among attributes value of new case with their respective attribute value in the old cases from the case base. The following are the local similarity functions used in this study

- **Equal:** Equal local similarity implies that the value of input attribute and attribute value of cases on the case base must be match.
- **Interval:** When you select interval similarity and adjust interval value. It is not mandatory to input match value, jCOLIBRI matches by keeping in mind the interval, any value can compromised along the interval value.

B. Global similarity measurement

The global similarity measurement allows computing the similarity between two cases namely an old case from the case base with new case from the user. This measurement computes the similarity by using all local similarity of attributes of the cases and it is used to select the most relevant cases. Average similarity is the global similarity used in this research.

- **Average:** Average similarity is a type of global similarity that considers the average of all attribute local similarity values to select relevant cases. The algorithm for average similarity works as follows [23,21,10].

Step 1: Find the local similarity of attributes of a new case for all attributes of the case which make up the case base

Step 2: Multiply the result of the local similarity of attributes with their corresponding attribute weight (importance value).

Step 3: Add the value of all attribute results of step 2.

Step 4: Add all weights of attributes that represent the importance value of the attributes and multiply by the number of attributes.

Step 5: Divide the result of step 3 by the result of step 4 and the result of this step is the global similarity that represents the degree of match of the old case with the new input case.

6. Designing the Cbrfmd

Integrating the components to provide the general case based reasoning framework for malaria diagnosis is the next task once the architecture and other required sub activities are done. To do this the components such as the most determinant parameters, architecture and the prototypical representation are integrated to design the workable artifact as shown in the figure 4. Below

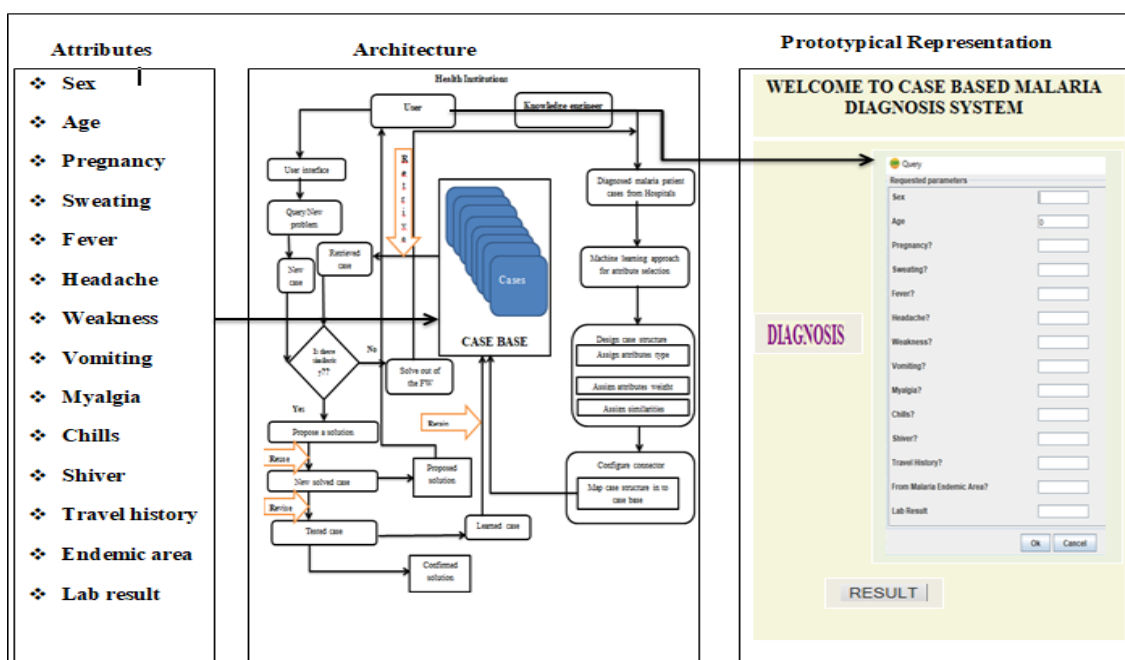


Fig.4. Integrated components of CBRFMD.

7. Testing and Evaluation

Testing and performance evaluation of the CBRFMD is also one from the objectives of the study. To attain this, performance measurement of the designed framework on its main CBR cycles (statistical evaluation), case similarity evaluation, comparative evaluation and user acceptance evaluations are main evaluation techniques used.

7.1. CBR Cycles (4Re) Evaluation of the CBRFMD by Using Statistical Analysis

Statistical evaluation is main evaluation technique in CBR framework. In this study it is used to evaluate the retrieval capability by taking individually all cases as a test case turn in turn (leave-one-out of cross validation testing proportion) and the whole case base as a training cases which comprises the case base. To do this both recall and precision are the two main statistical analysis used in this study to evaluate the retrieval capability of CBRFMD. The case base of this research contains seventy (70) cases due to this totally 70 tests could be conducted to come up with the retrieval performance of the framework.

A. Evaluation of Retrieval Cycle Performance

Precision and recall are the two statistical analysis measures used to evaluate retrieval performance and effectiveness of this framework. Precision is the ability of the framework to retrieve more relevant cases to a given query beside the specified threshold similarity boundary, whereas recall is the ability of the framework to retrieve all relevant cases to the current new problem [24,25,26]. Moreover, to incorporate the precision and recall evaluation the relevant malaria cases from the case base for each test cases should be known. This task is done by the domain experts. First the domain experts are given the test cases to assign possible relevant cases from the case base. To assign relevant cases for each test case the domain experts mainly use the value of solution attribute and their general experience. The main thing that should be given high emphasis in calculating precision and recall is the threshold interval value used, since it affects the value.

Table 4. Sample Test Cases and their Relative Cases Assigned by Domain Experts.

Test cases	Identified relevant cases from the case base
Case30	case2, case11, case16, case20, case24, case26, case63, case66
Case33	case4, case12, case13, case21, case27, case41, case43, case48, case54, case65, case68
Case34	case5, case17, case22, case28, case45, case51, case58, case70
Case35	case1, case8, case10, case18, case19, case23, case25, case29, case32, case39, case42, case46, case49, case52, case59, case67, case69
Case36	case6, case9, case15, case44, case50, case57, case62
Case37	case31, case37
Case38	case7, case38
Case39	case1, case8, case10, case18, case19, case23, case25, case29, case32, case35, case42, case46, case49, case52, case59, case67, case69
Case40	case3, case14, case47, case53, case60, case61, case64

Even though, there is no standard threshold interval stated for the degree of similarity in retrieving cases for any case based reasoning exhibition, different threshold intervals have been used by different researchers for their case similarity. In the domain of health both Henok[5] and Getachew [10] uses threshold interval of [1.0, 0.8). This implies that any case from the case base scored global similarity of 80% and greater than 80% with the queries are retrieved. As the threshold level closes to 1.0 the value of precision increase and the value of recall decreases, whereas the value of precision decrease and recall increases as threshold level goes far away from 1.0. Different threshold level case similarity is tested for this study by the researchers such as [1.0, 0.5), [1.0, 0.6), [1.0, 0.7) [1.0, 0.8) and [1.0, 0.9) to be selected and used. When the highest threshold level ([1.0, 0.9)) is

tested the precision is much better than the lower threshold level ([1.0, 0.5)) whereas the recall value is very low and the reverse is true as the lower threshold level is tested. Little reduction on precision also observed while [1.0, 0.6) and [1.0, 0.7) are tested. The threshold level case similarity [1.0, 0.8) obtains an average precision and recall when it is tested in its turn. After all selected threshold level case similarities are tested and analyzed by the researcher [1.0, 0.8) threshold level is used for this study. Finally with the threshold value of [1.0, 0.8) and leave-one-out cross validation testing proportion an experiment have been conducted for the whole cases (70 cases) to measure precision and recall performance of CBRFMD. But, due to time limitation only nine test cases are selected and confusion matrix is used to visualize the precision and recall measures.

Table 5. Recall and Precision Calculated for the Queries.

Test Cases	Retrieval performance measure				
	Total number of relevant cases	Relevant cases retrieved	Total number of retrieved cases	Recall	Precision
Case30	8	7	8	0.88	0.88
Case33	11	11	12	1.0	0.91
Case34	8	6	7	0.75	0.85
Case35	17	16	19	0.94	0.84
Case36	7	5	8	0.71	0.62
Case37	2	2	4	1.0	0.50
Case38	2	2	4	1.0	0.50
Case39	17	16	19	0.94	0.84
Case40	7	6	7	0.85	0.85
Total	79	71	89	0.89	0.79

B. Evaluation of Reuse Cycle

In any CBR approach exhibitions as well as in CBRFMD, after the relevant cases are retrieved to the current query at retrieval phase the next task is reusing the solution of the best selected case to solve the new problem or case. The first and most objective of reuse cycle is driving old solutions to the new problem (case) and solving it correctly. To evaluate reuse performance of this study an accuracy measurement is undertaken. Accuracy is the proportion of number of correctly solved cases to the total number of tested cases. To do this leave-one-out cross validation testing proportion is conducted by the researchers for 70 cases experimentation and an exciting accuracy is registered which is hopeful result as shown in the table 6 below.

Table 6. Reuse Cycle Accuracy Value

Total Leave-one-out cross validation testing experiment taken	Correctly diagnosed cases	Incorrectly diagnosed cases	Accuracy
70	64	6	91.4%

7.2. Evaluation of working with uncertainty and learning (Retaining) Capabilities of CBRFMD

Working with uncertainty and learning capabilities are the most advanced features of case based reasoning techniques which makes CBR more preferable than other reasoning techniques. The working with uncertainty capability enables the CBR frameworks to solve novel problems which have not perfect matched case in its case base by adapting solution from old similar cases whereas learning capability allows CBR frameworks to learn from these novel solved problems/cases/ and to use it for future problem approaching. Such capabilities of this framework are tested by obtaining new malaria case. This new malaria case (with its description and solution attribute) is first evaluated by the domain experts for its completeness and it is stated as follow:

A patient is 39 years old and she has pregnancy of seven months. She is from malaria endemic area (woreda Tsegedie, Tabiya Alemgentet, western Tigray zone). Sweating, high fever, feeling of weakness, headache for two consecutive days, vomiting and body shiver were the observed symptoms before laboratory test. After the laboratory examination, her blood film shows numerous and small ring trophozoites (RT). Finally, all assessment reveals she was being attacked by plasmodium falciparum parasite malaria type. The treatment recommended by the physician was; she has to swallow oral quinine of three from 200mg and two from 300mg salt three times per day for seven days consecutively.

To test the working with uncertainty and learning capability of this study, the problem description of the above novel case was fed to the CBRFMD. The CBRFMD retrieves relevant cases from the case base by computing the similarity of the case base with the query. Even though no case from the case base have global similarity of 1.0 with the new query, the framework proposes a solution and the revise cycle of the framework brings the new malaria case with its recommended solution for verification purpose as shown in the figure 5 below and this shows the working with uncertainty capability of the framework. The solution recommended for the new malaria case by the CBRFMD and physician was similar.

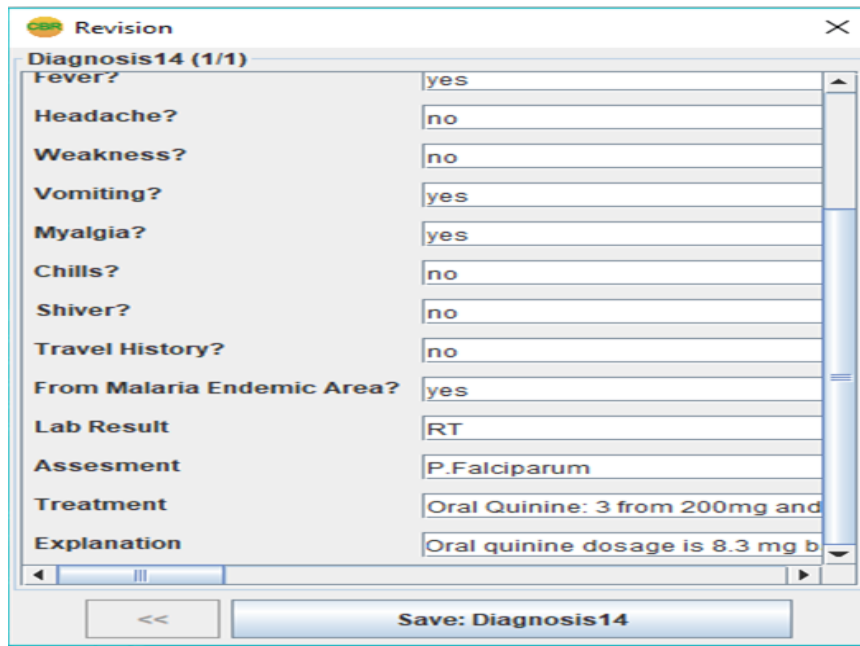


Fig.5. Working with uncertainty capability of CBRFMD and revision cycle

Finally after the new solved malaria case is validated the new case should be retained for future use in the case base as 71th case as shown in the figure 6 below and this also shows the learning capability of CBRFMD.

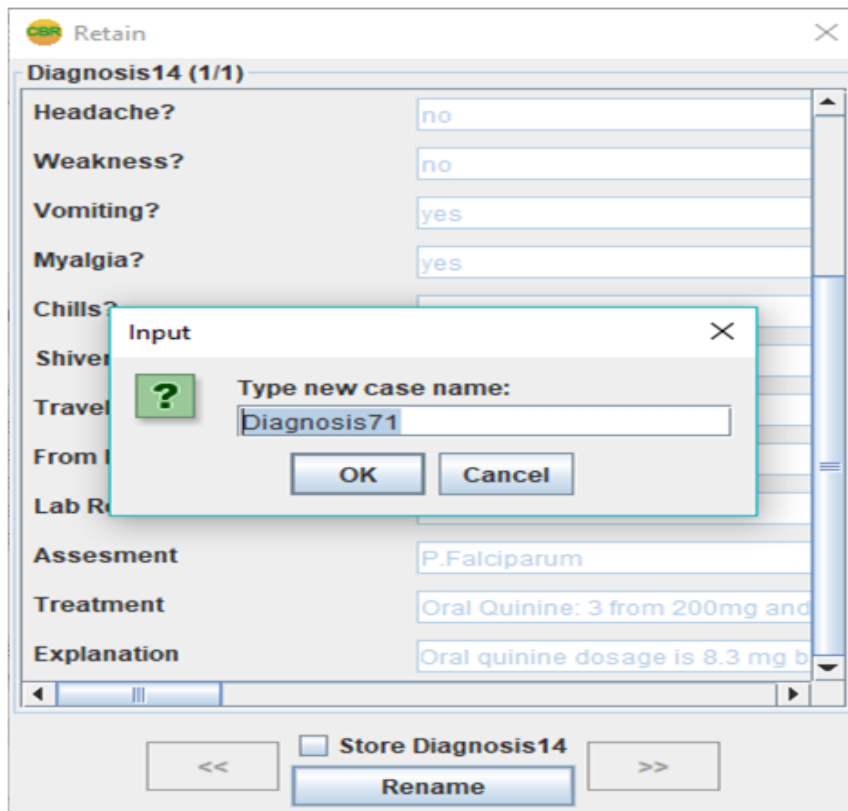


Fig.6. Retaining and learning of newly solved cases

7.3. User Evaluation of CBRFMD

The user acceptance testing for this research work was conducted in the real situation at Ayder specialized and comprehensive hospital to test and validate the CBRFMD, specially the selected attributes to represent malaria cases, applicability and the significance gained from this diagnosis framework. To do this eight health professionals are selected by the researchers purposely; two of them are physicians who have experience in malaria diagnosis and

involved in the research work as domain experts, three from health officer staffs, two from nurses and one from department of medicine graduating class student. Here, all the selected domain experts are guided by the researcher how they have to interact with CBRFMD. Then, malaria test cases' problem description is given to all selected experts to feed the test case in to the framework and to observe how the framework approach for solving the new malaria case as well as to interpret the result of the framework several times. At the end after the experts use practically and look the way CBRFMD attempt to solve the test case problems, they are asked to read and rank the evaluation parameters used in the evaluation criteria as: 1= poor, 2= fair, 3= good, 4= very good and 5= excellent.

Table 7. User Evaluation Results

Evaluation parameters	No. of respondents in each performance evaluation value					
	Poor	Fair	Good	Very good	Excellent	Average
Does the decision making is adequate and clear?	0	2	5	1	0	2.9
Are the retrieved cases are relevant in the decision making?	0	0	1	6	1	4.0
Does the proposed solution fits with the problem description at hand?	0	0	0	4	4	4.5
Is CBRFMD is easy to use, efficient in time and solves the regional problem in Tigray	0	1	1	4	2	3.9
Does the attributes representing the Malaria cases are relevant?	0	0	2	5	1	3.9
Is CBRFMD is applicable for Malaria diagnosis?	0	0	1	3	4	4.4
Is the explanation given by CBRFMD is enough and complete?	0	0	0	8	0	4.0
	Total average					3.94

Generally, the evaluation result given by the domain experts to the framework is greater than an average for each evaluation parameter and the total average result is encouraging. Even though all the domain experts rank the evaluation parameter "Is the explanation given by CBRFMD is enough and complete" as very good they provide a feedback regarding the explanation capability of CBRFMD to include explanation about the side effects of medication and other supportive explanatory descriptions in the explanation facility which allows and supports users during adaptation phase.

7.4. Case Similarity testing

The main objective of this test was to know the efficacy of CBRFMD in selecting the appropriate malaria cases for the new malaria case or query. To attain this, the researchers conduct an experiment with three different experiment groups. Cases with their least weighted attribute's values modified forms the first experiment group, whereas the second experiment group contains cases with their highest weighted attribute's value is modified and the third experiment group are cases from the case base without any modification of their attribute's value. The case similarity test is performed by presenting each test case in the CBRFMD individually and the selected cases are evaluated against the expected selection to measure the performance. The following are sample test cases presented to the CBRFMD for case similarity measurement purpose.

- **Experiment one problem description:** a case it's least weighted attribute value is modified.

A 27 years old pregnant woman patient who came from malaria endemic area and having symptoms of high fever, feeling of weakness, headache for more than a week and body shiver. In addition to the identified symptoms the laboratory result of her blood film indicates the availability of ameboid trophozite(AT) parasites in the red blood cells.

- **Experiment two problem description:** a case it's least weighted attribute value is modified.

A 70 years old woman, who lives for more than 30 years in malaria endemic area up to recent, was with symptoms of high fever, sweating and chills. After these symptoms are identified, her laboratory blood film result also shows the availability of large rounded gametocycle(LRG) parasites within the red blood cells.

- **Experiment three problem description:** a case it's highest weighted attribute value is modified

A 18 years old pregnant woman patient who came from malaria endemic area and having symptoms of no fever, feeling of weakness, headache for more than a week and body shiver. In addition to the identified symptoms the laboratory result of her blood film indicates the availability of ameboid trophozite(AT) parasites in the red blood cells.

- **Experiment four problem description:** a case it's highest weighted attribute value is modified:

An old man aged 70 years old, who lives for more than 30 years in malaria endemic area up to recent, was with symptoms of high fever and chills.

After these symptoms are identified, his laboratory blood film result also shows the availability of large rounded gametocycle(LRG) parasites within the red blood cells.

- **Experiment five problem description:** a case with no attribute value is modified:

A 18 years old pregnant woman patient who came from malaria endemic area and having symptoms of high fever, feeling of weakness, headache for more than a week and body shiver. In addition to the identified symptoms the laboratory result of her blood film indicates the availability of ameboid trophozite(AT) parasites in the red blood cells.

- **Experiment six problem description:** a case with no attribute value is modified

An old man aged 70 years old, who lives for more than 30 years in malaria endemic area up to recent, was with symptoms of high fever, sweating, chills. After these symptoms are identified, his laboratory blood film result also shows the availability of large rounded gametocyte(LRG) parasites with red blood cells.

Table 8. Sample Queries Used for Case Similarity Measurement Experimentation

Test cases	Problem description	Expected similar case in the case base	Altered attribute of cases
Query1	Experiment1	Case14	The value of attribute age is modified
Query2	Experiment2	Case29	The value of attribute sex is modified
Query3	Experiment3	Case14	The value of attribute fever is modified
Query4	Experiment4	Case29	The value of attribute sweating is modified
Query5	Experiment5	Case14	No attribute value is modified
Query6	Experiment6	Case29	No attribute value is modified

Once the queries and expected selection are known, each test case is presented to the CBRFMD and the following case similarity between test cases/queries/ and their respective cases in the case base is observed as depicted in the table 9 below.

Table 9. Case Similarity Test Measure

Queries	Global similarity with expected cases	Expected case in the case base
Query1	0.97	Case14
Query2	0.94	Case29
Query3	0.89	Case14
Query4	0.82	Case29
Query5	1.0	Case14
Query6	1.0	Case29

This case similarity test experiment result reveals that when a query has similar attributes with a case stored in the case base, the global similarity is scored for 1.0 and it is called exact match as query5 and query6 of table8. Whereas the global similarity decreases as the attribute values are changed. The weight of attribute changed greatly affects the global similarity. As highest weighted attribute value is changed the global similarity is highly affected than least weighted attribute value is changed as query1 and query2 of table 9. Thus the framework achieves highest case similarity to support users in being with most similar malaria cases for queries.

7.5. Comparison of CBRFMD with previous CBR systems

In the domain of medicine researchers applied case based reasoning for different areas in Ethiopia. The performance of these systems had been evaluated by the respective researchers and their overall result is compared against CBRFMD as depicted in the table 10 below.

Table 10. Comparison of CBRFMD with Previous Works in Medical Domain

Specific research area	Resources in the cases base	tool	Case attribute selection	Retrieval performance		Reuse performance
				Recall	Precision	Accuracy
AIDS Alemu[50]	50cases	JCOLIBRI	Manually	72%	63%	Not evaluated
Hypertension Henok[8]	45cases	Python	Manually	86.1%	60%	88.89%
Anxiety Getachew[9]	50cases	JCOLIBRI	Manually	82%	71%	Not evaluated
Malaria	70cases	JCOLIBRI And Eclipse	Machine learning approach used. information gain algorithm	89%	79%	91.4%

The main reason for the better result value of recall and precision registered for this study is the number of cases used. Because, as the number of cases increased the performance of the framework is also increased due to wider probability of retrieving the most relevant cases from the case base. As depicted in the table 10 above the total number

of cases used in this study is increased by around 50% of the previous studies number of cases and this is the main reason for the higher performance result value achieved in this study. Beside this the integration of machine learning approach in this study for attribute selection took the higher role for the higher accuracy of results registered by CBRFMD, because attribute selection leads to removal of irrelevant attributes which might be the potential cause of inaccuracy or less accurate in any framework.

8. Discussion of Results

Evaluating the designed framework in different mechanisms was one objective of this study and different evaluation techniques have been applied to measure the performance of the framework such as statistical analysis, user acceptance testing, case similarity testing, an experiment and comparative evaluation. The result gained from this study is interesting and it is discussed against the research questions stated in the problem statement section of this research. In addition to achieving the objective, this study stands with five main research questions such as discussed as follow.

What are the most determinant symptoms and parameters in the malaria diagnosis framework: were the first and second research questions of this study respectively.

In the process of malaria diagnosis symptoms and related factors identified in the patients are the main inputs considered in diagnosing malaria suspected patients. The degree of their relevance in the diagnosis varies and still there is no scientific evidence to know and rank the degree of relevancy of symptoms and other factors. This limitation could also lead to inaccurate diagnosis which could be a cause for loss of human life. But identifying and knowing the most determinant symptoms and parameters to be considered during malaria diagnosis helps practitioners to make more emphasis on these parameters while diagnosing malaria patients. To overcome these problems information gain machine learning approach is used in this study to select the most relevant symptoms and parameters on the collected historical patient cases. To this end the most determinant symptoms and parameters in the malaria diagnosis are identified. Thus, age, sex, pregnancy, sweating, fever, headache, weakness, myalgia, vomiting, chills, shiver, travel history, from malaria endemic area and lab result are the most determinant symptoms and parameters whereas anemia and loss of appetite are found to be less relevant symptoms from the malaria patient cases. Finally CBRFMD is designed by using the identified most determinant symptoms and parameters to assist practitioners in diagnosing malaria easily.

On the other hand the results gained from the performance evaluation of this study are promising. In addition to this the user acceptance testing conducted on the domain area also scores more than an average result. Specifically for the evaluation parameter of user acceptance testing "*Is CBRFMD is applicable for Malaria diagnosis*", 50%, 37.5% and 12.5% of domain experts rate the designed framework as excellent, very good and good respectively. This rating of respondent shows the framework could be considered as a remedy in Tigray health institutions and it could improve the service of malaria diagnosis. Beside to this "*Is CBRFMD is easy to use, efficient in time and solves the regional problem*" is also another evaluation parameter in the user acceptance testing and it was rated as excellent, very good, good and fair by the 25%, 50%, 12.5% and 12.5% of domain experts respectively. This rating percentage by the domain expert shows the problem of inexperienced health professionals can be solved by the implementation of this designed framework in rural areas. Therefore, CBRFMD could acceptable in Tigray health institutions and it could be good remedy to solve lack of experienced practitioners in rural areas of the region. Generally, case based reasoning framework for malaria diagnosis are solutions for medicine domain mainly for malaria disease detection and diagnosis easily and 88.9% an average accuracy registered from CBRFMD could be an evidence for case based reasoning methodology promising performance for this domain.

9. Conclusion and Futurework

Now a day providing early and quick diagnosis service for patients in Tigray is difficult task with the current manual system. To overcome such problems medical diagnostic systems should implemented to provide quick diagnosis service for patients and support practitioners in decision makings. In this study we design CBR diagnostic framework for malaria diagnosis to minimize the overall challenges in malaria diagnosis process. The framework is designed by identifying the most relevant factors in malaria diagnosis using machine learning approach and its performance is evaluated by the domain experts in the actual health institutions by using purposive sampling technique. As a result 78.8% of the experts accepted it. Academically this paper contributes growing of demands of using medical diagnosis systems which employs recent artificial intelligence techniques like CBR in health institutions. At the end, the future works we forwarded are an investigation to keep the highest retrieval performance and maintain the case base as the size of the case base grow, developing hybrid framework to include best features of rule based reasoning technique and localize the framework for self-diagnosing with mobile phones.

Acknowledgment

We would like to acknowledge hart fully to NUMB-MU for the support of 20,000 NOK used for successfully finish this study.

References

- [1] Eliyas N. "Prevalence of malaria and its biomedical knowledge among households in ayira district, western Ethiopia". Unpublished Master's Thesis, Haramaya University, Ethiopia. 2014
- [2] Chala D., Million M., Debela T, "Developing a Learning Knowledge-Based System For Diagnosis And Treatment Of Malaria", IJCSI International Journal of Computer Science Issues, Volume 13, Issue 4, July 2016.
- [3] Aster M. "medical expert system in tigray health institutions: the case of malaria" Unpublished Master's Thesis, Axum University, Ethiopia 2016.
- [4] Shortliffe, H.E, "Computer-Based Medical Consultations": MYCIN, New York: Elsevier/North Holland Publishing Company, 1976.
- [5] Getachew W. "Application of Case-Based Reasoning for Anxiety disorder diagnosis. Unpublished Master's Thesis, Addis Ababa University, Ethiopia, 2012.
- [6] Ahmed M., Begum S., Funk P., Xiong N., Scheele A. "Case-Based Reasoning for Diagnosis of Stress Using Enhanced Cosine and Fuzzy Similarity". *Transactions on Reasoning for Multimedia Data*, 1(1), 3-19. 2008.
- [7] "Annual reports of five years (2004-2008) of all diseases reports in one" ,TRHB (Tigray Regional Health Bureau) Mekelle, Tigray, Ethiopia, July, 2017
- [8] Alemu J. "A Case-Based Approach for Designing Knowledge Base System for Addis Resource Center (ARC):", The Case of Warmline Clinician Consultation Service, Unpublished Masters' Thesis, Addis Ababa University, Ethiopia. 2010.
- [9] Henok B. "A Case-Based Reasoning Knowledge Based System for Hypertension Management". Unpublished Master's Thesis, Addis Ababa University, Ethiopia. 2011.
- [10] Koton, P. (1988). Reasoning about Evidence in Causal Explanations. In the Proceedings of the American Association for Artificial Intelligence, Morgan Kaufmann, San Mateo, CA, pp. 256-261.
- [11] Bichindaritz, I. & Seroussi, B. (1992). Concept Learning in a Real World Domain. *Technique et Sciences Informatiques*, 11 (4), 69-98.
- [12] Bichindaritz, I. (1995). Case-Based Reasoning Adaptive to Several Cognitive Tasks. Paper presented in the Proceedings of the International Conference on CBR, LNAI, No. 1010, Springer, Berlinpp, pp.391-400.
- [13] Bichindaritz, I., Kansu, E. & Sullivan KM. (1998). Case-Based Reasoning in CARE PARTNER: Gathering Evidence for Evidence-Based Medical Practice. In the Proceedings of the 4th European Workshop on CBR, Smyth B & Cunningham P, (Eds.), Springer, Berlin, pp. 334-345.
- [14] Biazen G. "Application of case based recommender system in field of study selection." Unpublished MSc. Thesis, Addis Ababa University, Ethiopia. 2013.
- [15] Yibeltal, C. "Application of recommender system in investment sector", MSc Thesis, School of Information Science, Addis Ababa University, Ethiopia. 2013.
- [16] "Federal democratic republic of Ethiopia ministry of health", National malaria guidelines 3rd edition, Addis Ababa. January 2012.
- [17] Makhfi "Introduction to Knowledge Modeling and Neural Network." Available at http://www.makhfi.com/KCM_intro.htm accessed date [Accessed October 1, 2018].
- [18] Fag H, et al. "Case Based Reasoning for Logistics Outsourcing Risk Assessment Model", Proceeding of International Conference on Enterprise and Management Innovation, pp. 1133-1138. 2007.
- [19] Stahl, A., & Roth-Berghofer, T. "Rapid Prototyping of CBR Applications with the Open Source Tool myCBR", in *Advances in Case-Based Reasoning*, 615-629. 2008.
- [20] Salem, M., Rooushdy, M. & Hodhod, A. "A Case Base Experts System for Diagnosis of Heart Disease". *International Journal on Artificial Intelligence and Machine Learning*, 5(1), 33-39. 2005.
- [21] Lützel Schwab, S. "Case-based Reasoner for OWL-S Web Services:" An Experiment and Task Perspective, Unpublished Diploma Thesis, Dynamic and Distributed Information Systems, Kaiseraugst AG, University of Zurich, Switzerland. 2007.
- [22] Watson, I. & Marir, F. "Case-Based Reasoning:" a Review. *The Knowledge Engineering Review*, vol. 9, issue 4, pp 327-354, 1994.
- [23] Junker, M., Hoch, R., & Dengel, A. "On the Evaluation of Document Analysis Components by Recall, Precision and Accuracy", *Proceeding of the fifth International Conference on Document Analysis and Recognition*, Berlin, pp. 713-716. 1999.
- [24] Losee, R.M. "When Information Retrieval Measures agree about the Relative Quality of Document Rankings", *Journal of the American Society for Information Science*, vol. 51, pp. 834-840, 2000.
- [25] McSherry, D. "Precision and Recall in Interactive Case-Based Reasoning", In *Case-Based Reasoning Research and Development (ICCBR)*, Lecture Notes in Artificial Intelligence vol. 2080, pp. 392-406, 2001.

Authors' Profiles



Mr. **Eshetie Gizachew Addisu** is currently working as lecturer in school of computing, Ethiopian Institute of Technology-Mekelle, Mekelle University Ethiopia. He obtained his MS.c. Degree in computer science from Mekelle University, Ethiopia and BS.c. Degree in Information Systems from Debre berhan University, Ethiopia. His research area of interest is artificial intelligence and machine learning.



Dr. **Abiot Sinamo Boltena** is currently working as a Director General, ICT sector at Federal Democratic Republic of Ethiopia, Ministry of Innovation and Technology. He obtained his PhD. in Intelligent systems and ERP systems from Oldenburg University, Germany, and MSc. Degree in Information Science from Addis Ababa University, Ethiopia, and BSc. Degree in Construction Technology from Nazereth Technic College, Ethiopia. His research area of interest is artificial intelligence and robotics.



Mr. **Samson Yohannes Amare** is currently working as lecturer in school of computing and chair head of software engineering chair, Ethiopian Institute of Technology-Mekelle, Mekelle University, Ethiopia. He obtained his MS.c. Degree in Informatics Engineering from University of Algarve, Portugal, and BS.c. Degree in Information Systems from Addis Ababa University, Ethiopia. His research area of interest includes digital health, Block chain, data science, image processing and retrieval.

How to cite this paper: Eshetie Gizachew Addisu, Abiot Sinamo Boltena, Samson Yohannes Amare, "Case-Based Reasoning Framework for Malaria Diagnosis", *International Journal of Information Technology and Computer Science(IJTCS)*, Vol.12, No.6, pp.31-48, 2020. DOI: 10.5815/ijitcs.2020.06.04