Asian Journal of Advances in Research

4(1): 736-749, 2021

DEVELOPMENT OF A RULE-BASED EXPERT SYSTEM FOR THE IDENTIFICATION OF COMMON TERRESTRIAL AND AQUATIC PLANTS

DIPO THEOPHILUS AKOMOLAFE[1*]

[1]Department of Mathematical Sciences, Olusegun Agagu University of Science and Technology, Okitipupa, Nigeria.

AUTHOR'S CONTRIBUTION

The sole author designed, analysed, interpreted and prepared the manuscript.

Received: 25 May 2021
Accepted: 31 July 2021
Published: 04 August 2021

Original Research Article

ABSTRACT

The advancement in technology in this era has given rise to the development of many expert systems to substitute and aid experts and non-experts in the accomplishments of various tasks. This, however, gave birth to this research as it stands out in the design and development of an expert system to help in the identification of some common terrestrial and aquatic plants to aid users in gaining and affirming taxonomic and botanical origins in the identification process. The design of this expert system takes advantage of architecture for the construction of expert systems in the rule-based domain. The information gathered in the heuristic and tacit form was analysed and operated upon based on the taxonomical morphology used in the identification of some common\selected plants and its certainty score and then turned to rules for easy programming.

The system was developed using Microsoft .Net Framework 4.5.2 (C#, Telerik framework, Chen0040 cs expert shell Nugget, Aforge Framework) due to their robustness and large support group. The system was successfully tested and evaluated using real life data.

Keywords: *Expert system; leghaemoglobin; haemoglobin; inference engine; knowledge base; haemoprotein; holoprotein.*

1. INTRODUCTION

An examination into the individual balance diet needs which consists of Carbohydrates, Fats, Proteins, Vitamins, Fibre has shown that protein unlike carbohydrates and fats, are mostly sourced from animals (Red meat) which are in limited quantity and leads to many health issues as discussed in Sammy [1] and Mark [2], while similar source (sustainable) of protein can be got from leguminous plants (Leghaemoglobin) without the side effects of consuming red meats [3,4].

Many organizations have shown great interest in using information technology in reducing the complexity and iterative processes of educating, training and sharing expert knowledge on leghaemoglobin without the increased lead times. The proposed system has been subjected to quantitative data collected from botanist, biochemists and published papers. The data was relatively modified into an expert system knowledge base to help in various information needs on leghaemoglobin.

This system was developed, based on a type of simulation which requires no chemical or laboratory procedure in the identification process of symbiotic leghaemoglobin in plants root nodules, but can identify symbiotic leghaemoglobin via a morphological approach/procedure which are

**Corresponding author: Email: dtakomolafe@yahoo.com*

formulated into rules; rules which consists of antecedents leading to a viable accuracy rate as equivalent to a human expert. Leghaemoglobin is a nitrogen or oxygen carrier and haemoprotein is found in the nitrogen-fixing root nodules of leguminous plants (e.g. Soybean, Cowpea, Alfalfa). Leghaemoglobin has close chemical and structural similarities to haemoglobin, and, like haemoglobin, is red in colour. The holoprotein (protein + heam cofactor) is widely believed to be a product of both plant and the bacterium in which the apoprotein is produced by the plant and the heam (an iron atom bound in a porphyrin ring) is produced by the bacterium as stated in Mark [2].

Many research works had been carried out in the area of intelligent systems, for solving different problems but enough has not been done in terms of developing an intelligent system on leghaemoglobin most especially in the botanical and biochemical origins of plant haemoglobin. It is against this backdrop that this research was carried out.

2. RESEARCH OBJECTIVES

The primary objective of this research is to develop an Expert System that can be used to get vital information and answer important questions about leghaemoglobin, while the specific objectives are to:

i. generate a multidimensional overview of the structure, function, and importance of leghaemoglobin;
ii. manage and develop a dynamic system for expert information exchange and
iii. showing/enlighten users on basic utilization of leghaemoglobin.

3. RELEVANT WORKS

There are many areas in biochemistry and botany where an expert system has been designed and implemented to provided solutions and quality information to users based on the specific domain in question. Among these diverse domains include an expert system for identification of a sample of Sponges, an identification of industrial location factor analysis and a self-learning computer troubleshooting system. Domingo [5] in his research work for the expert system spongia, used MILORD II which is a form of Common Lisp and runs on a Macintosh environment only, knowledge contained in the knowledge base are structured in compartments called modules, which allow an incremental approach to the knowledge base of the ES, his expert system works by comparing user data to corresponding rules in the knowledge base and performing a best fit inference to the data, thereby producing an inference on likely sample sponge of user. Also Amanuel [6] followed a rule-based approach to develop a system that helps in identifying problems by asking users a list of guiding question and when answered, provides the system with information about what problem Sets the issue may belong to, and which one it does not belong to and then generate another list of questions that apply to subsets of those identified questions.

Furthermore, Aleksander [7], Ayodeji [8], Giarratano [9] and Hatzilygeroudis [10] followed a fuzzy logic approach for meeting end-user queries. The system works in identifying various factors (e.g. transportation, labour, raw materials, markets, utilities and etc) which contributes to the success of an industry in a particular location.The system manages and explores the knowledge in the application by reasoning on a database of facts by means of suitable inference rules. Fuzzy logic is applied to the decision support system for factors analysis providing a formal methodology to capture valid patterns for uncertainty reasoning. The homogeneous framework uses a set of Fuzzy Inference Systems to interpret, standardize and fuse heterogeneous data to estimate normalized industrial location factors Nureize [11] Patra [12] Cuena [13]. Yerokun and Onyesolu [14] developed a neuro-fuzzy expert system for detection of leghemoglobin in legumes, evaluated and knowledge acquisition was done by oral interview of prominent biochemists and botanists that provided key technical facts on leghemoglobin and visits to botanical gardens of Society for Underutilized Legumes (SUL) in Nigeria. In the development, production rule-base technique and forward-chaining mechanisms with linguistic antecedent conditions were used. MATLAB platform was employed for the development of the system. Confusion matrix was employed for the performance evaluation of the developed system. The result is a neuro-fuzzy expert system with gaussian membership functions with accuracy of 100% as against 99.56% for triangular, trapezoidal and gaussian combination functions, precision of 100% for all the membership functions evaluated and recall of 100% for gaussian membership functions and 99.53% for triangular, trapezoidal and Gaussian combination functions.

4. METHODOLOGY

4.1 System Architecture

This describes the fundamental components and knowledge for building the system into an expert system. The basic function of the system starts from the input which includes the user and knowledge engineer, works in the sense in which the knowledge engineer starts by adding or removing knowledge from/to the knowledge base by invoking some embedded ADO.NET code snippets, which is shown in the diagram below (Fig. 1). Here also, the user interacts with the user interface (interviewer components) with the aim of asking a query, the user interface contacts the inference engine for query processing, with either the inference giving a conclusion from the knowledge base or further interacting with users to give a more definite conclusion with previous queries held in the working memory. The inference engine conclusion solution process is also explained via the user interface (explanation Component). The knowledge base on the other hand interacts with various database to form a unified base containing the media contents which are hosted on the SQL server and the rules in text format (XML) which are updated by the knowledge engineer after gathering information (tacit and heuristic) from botanist, biochemist and published papers and transforming such data rules.

4.2 Components of the Architecture

The components of the architecture of the leghaemoglobin expert system is made up of

i. knowledge base;
ii. inference engine;
iii. the working memory;
iv. user interface and
v. Database

This followed the procedures stated in Jackson [15] and Samy [1]

4.3 Knowledge Base

The knowledge base of this system contains various morphological characteristics and common practices by botanist and biochemist in the identification of legumes plants, Rhizomes infected root nodules, and leghaemoglobin in root nodules of plants. The design of the knowledge base is distributed into Knowledge acquisition and Knowledge engineering.

4.4 Knowledge Acquisition

The knowledge acquisition component, in general simply refers to collation of knowledge from experts, knowledge gotten from experts was based on oral interview with prominent biochemist and botanist who helped in providing key technical facts about the subject domain leghaemoglobin and other reliable sources which includes certified published journals gotten off various servers on the internet and the orderly placement of such knowledge after conversion from the intermediate representation into an executable form (production rules) for the knowledge base.

The main means of interaction with the various experts was through an interview. Various tacit (unconscious knowledge), procedural (knowledge of how to do something) and declarative (knowledge of something being true or false) were captured via interview.

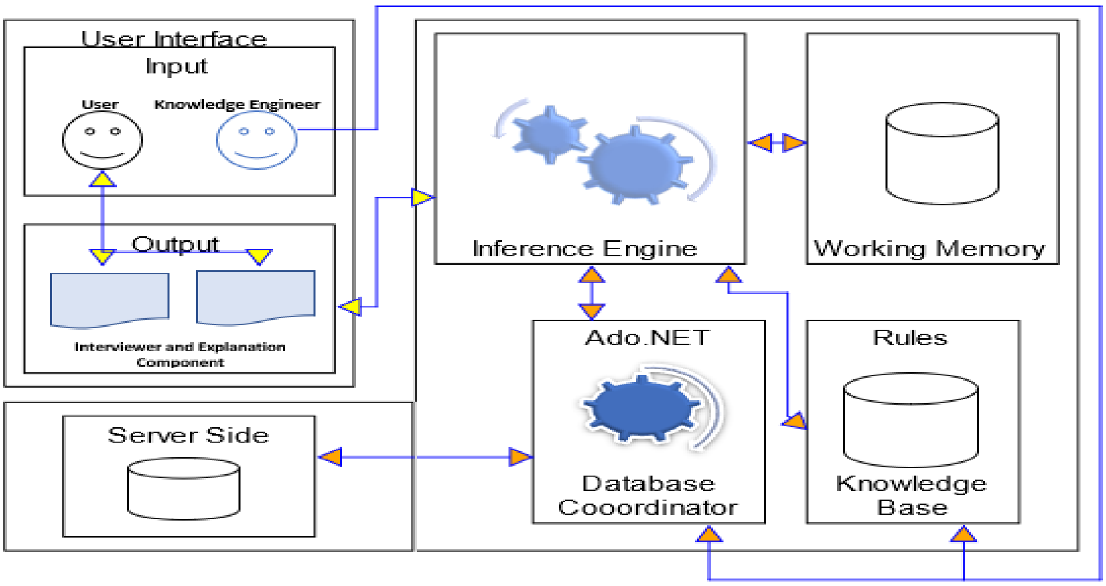

Fig. 1. System architecture

Interviewer & explanation Component

Some of the most important aspects of the subject domain (leghaemoglobin) explained by the various experts and implemented in the knowledge base of the expert system developed includes:

i. Identification of legumes from other plants
ii. Identification of Rhizobia growth in nodules
iii. Identification of Symbiotic leghaemoglobin in root nodules of legumes
iv. The Extraction process of leghaemoglobin from Legumes via manual and automatic processes.

4.5 Knowledge Engineering

On completion of knowledge elicitation from the domain experts, the proceeding step is designing the knowledge base. In knowledge engineering, this involves picking the best method of representing the knowledge and the rules governing the inference making.

Class and objects were used to model the information elicited. A class named Rule was created and houses various objects such as: Rule(), AddAntecedent(), setConsequent(), AddRule(), AddFact(),and etc which serves as subsidiary methods. Each rule is saved into the system using the following object to the class Rule, some rules added contains media contents which are saved on the database, pointers are used to point to contents in the database using ADO.NET and Linq as an access control to the database (Microsoft SQL Server is the Relational database management system used in implementation).

4.5.1The inference engine

This is the brain of the expert system which helps in reasoning and conclusion derivation base on production rules in the knowledge base, it's also a subsystem that is directed through knowledge base rules to manipulate the knowledge in the working memory to draw an inference from it. The design and development of the inference engine takes advantage of chen0040 cs expert system shell.

The inference engine works by comparing facts supplied by users to system with the production rules already available in knowledge base. The rules occur in IF (antecedents), THEN (setconsequent). This mean if facts supplied by user matches a or sets of antecedents of rule then the consequent of such rule is fired by the inference engine which activates various actions by the system. Antecedents are expressions involving attributes connector "AND" in the inference engine. A sample of the rules used in the system for the identification of leghaemoglobin

```
rule = new Rule("Presence of Leghaemoglobin");
rule.AddAntecedent(new IsClause("Is Plant Leguminous", "yes"));
rule.AddAntecedent(new IsClause("Presence of Infection by Rhizobia", "yes"));
rule.AddAntecedent(new IsClause("Colour of traverse section of cutting of excised nodules", "milky"));
rule.setConsequent(new IsClause("Presence of Identification of Symbiotic Leghaemoglobin", "Positive certainty = 20%"));
rie.AddRule(rule);
```

According to the rules above which confirms with Garcia [16], Hustinawaty [17] and [18] the inference engine works using backward chaining, that is backward chaining in the sense that all rules antecedent is satisfied by the facts supplied by the user to the inference engine before the consequents are fired by the inference engine, affirming the presence of the molecule of leghaemoglobin in plant roots nodule.

In the system, the rule base of the inference engine is stated as follows:

Facts:	$F_1, F_2, \ldots, F_{n1}$
Antecedents:	$A_1, A_2, ..., A_{n2}$
Rules:	$R_1, R_2, \ldots, R_{n3}$
Consequents	$C_1, C_2, \ldots, C_{n4}$

The system only executes a consequent when a rule containing it is fired and a rule is executed only when all of its antecedents are satisfied by facts supplied to the user to the inference engine.

4.5.2 The working memory

It holds the data that is received from the user during the expert system session. Values in working memory are used to evaluate antecedents in the knowledge base. Consequences from rules in the knowledge base may create new values in working memory, update old values, or removeexistingvalues.

(en.wikibooks.org/wiki/Expert_Systems/Components_of_Expert_Systems).

4.5.3 The user interface

The user interfaces consist of graphical clickable controls which allow ease of input of queries by users and a natural language response controls to enable ease of use and understanding by expert and non-expert user.

Table 1. Knowledge for Identification Process base the on the following essential attributes Elicited from the Domain Experts

Identification of legumes from other plants						
Flower Petal Colour	Leaf Type	Leaf arrangement	Leaf blade edges	Leaf tip	Shape of leaf	Leaflet Number
Leaf Petiole presence	Flower symmetry	Flower Orientation	Fruit Type	Fruit Shape	Inflorescence	Inflorescence length
Hair on fruit	Plant Colour	Stem succulence	Stem Hairiness	Fruit Length	Seed Surface	Colour of Seed
Seed number	Stamen number					
Identification of Rhizobia growth in nodules						
Plant type	Presence of Visible nodules	Nodules on primary root	Nodules on lateral root	Presence of rhizome	Colour of traverse section of cutting of excised nodules	
Identification of Symbiotic leghaemoglobin in root nodules of legumes						
Plant type	Presence of infection by rhizobia		Colour of traverse section of cutting of excised nodules			

The user interface contains the following controls

i. Navigation controls (for ease of navigation between forms)
ii. Button controls (for executing of queries)
iii. Explanation controls (for explaining consequents of the inference engine)
iv. Media controls (for displaying of media contents from the database to users)

The user interface also consists of two basic components

i. The interviewer component: this interface is used for asking various questions from the user to gather facts for the inference engine to compare against the rules in the knowledge base of the system.
ii. The explanation Component: Conclusion/consequents are displayed in this component to users and antecedents leading up to the consequent are shown to the user from this component, reassuring users of the reasoning process. For example, it might output the rules used to come to the consequent.

5. RESEARCH FLOW

This point at the basic steps that were followed from the very beginning to the end of the research. The diagram below (Fig. 2) depicts that the research started by the consultation of multiple human experts and various published papers by experts on the subject domain leghaemoglobin and elicitation of useful information from them and ended at the conclusion and finalizing of the whole research work by documenting and tiding up its environment. The research begins with the consultation of experts and move to users' awareness and also to data coalition/analysis followed up to the transformation of such data in rules. The user interface follows in the process and then the system main components coding which includes the inference engine, etc., before the final testing and conclusion.

6. SYSTEM DESIGN PHASES

This comprises all the forms and phases that make up the whole system. Which include the welcome Form, which moves forwards to the main-form.

The main form consists of the following forms and phases with their sub-forms as listed below;

6.1 Hardware and Software Requirements

Expert System for detection of leghaemoglobin requires a system that has any one of Microsoft's Windows OS, Apple's Macintosh or Linux Debiandistro with either .Net framework 4.5 and above or Mono install. The hardware requirements include a PC with at least 512MB RAM and 500MB Disk space.

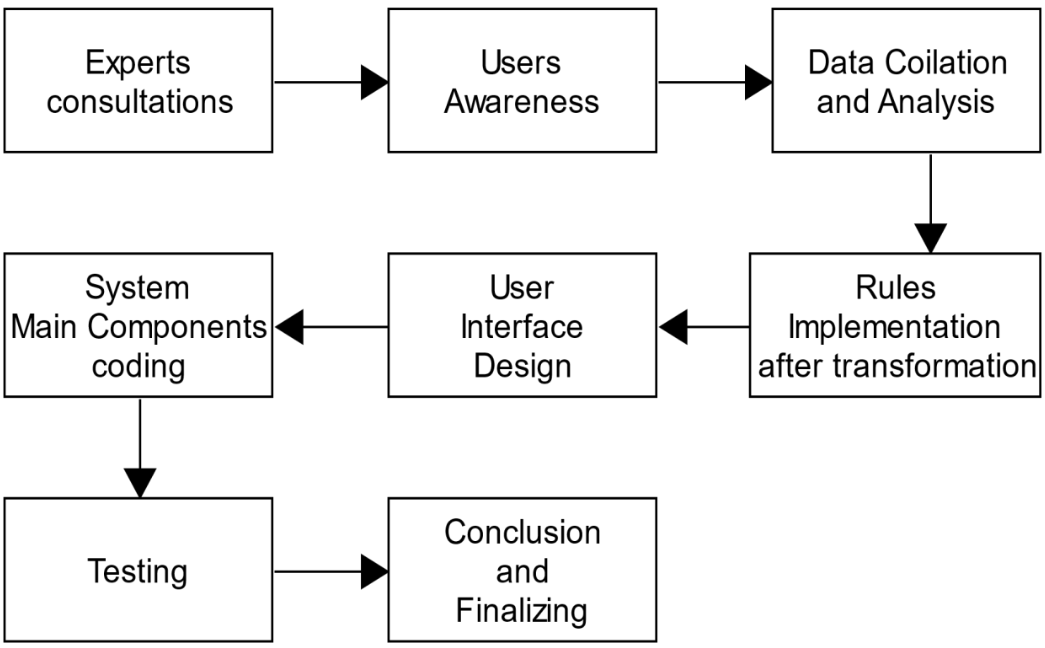

Fig. 2. Research flow

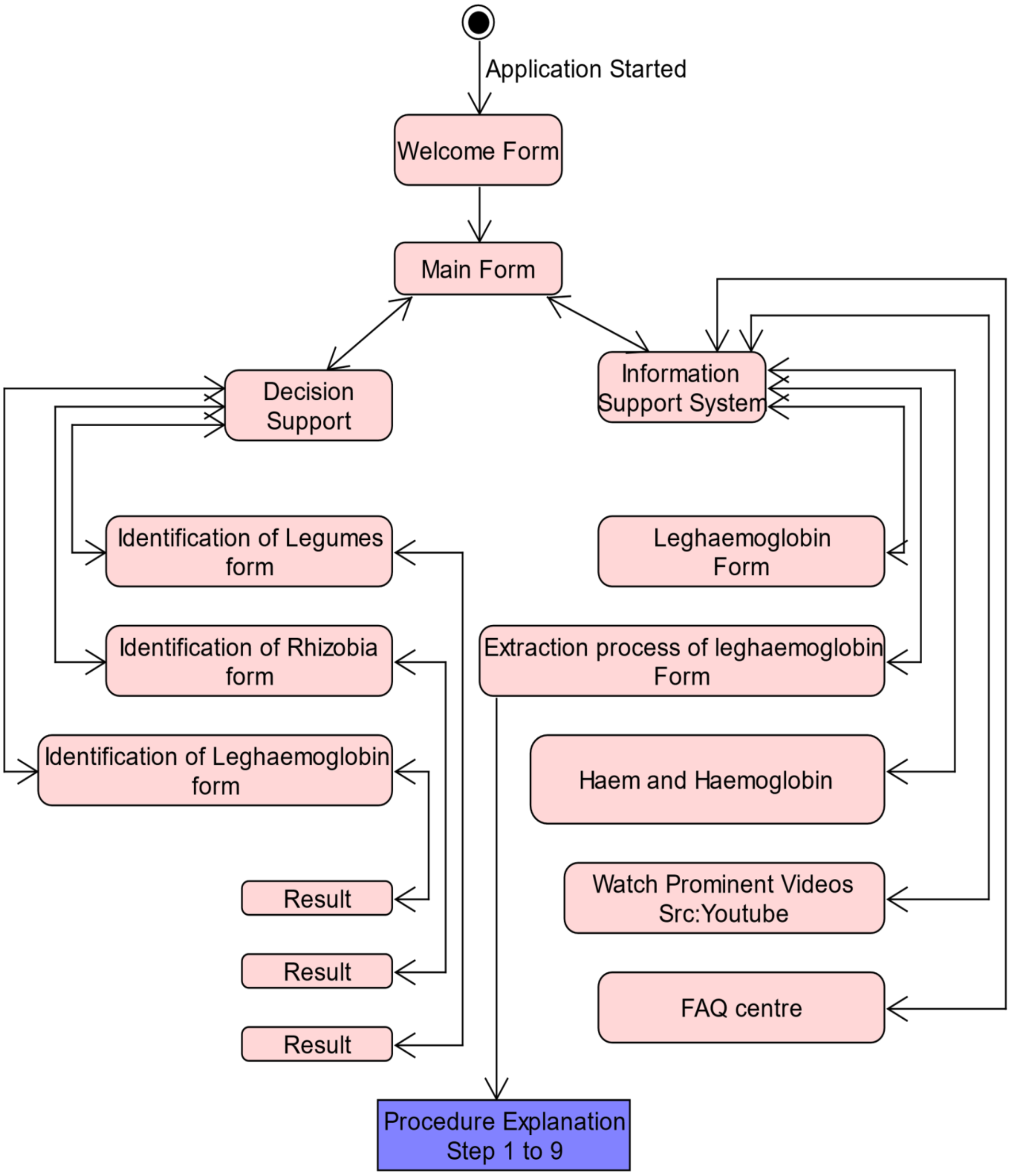

Fig. 3. State machine diagram of the system

6.2 Welcome Form

This is the first form displayed to the user on the launch of the software to the user via shortcuts created on the desktop or main program file. This form is represented in Fig. 5.

6.3 The Mainform

This is the next form displayed to the user to make a choice, either choosing between the decision support panel and Information Support. This form is represented in Fig. 6.

6.4 The Decision Support Panel

This form houses the expert decision-making inference engine which will assist a user in identifying a leguminous plant, Rhizobia in root nodules and leghaemoglobin molecules in plants by morphological means using either one of the interaction process incorporated into the system. This form is represented in Fig. 7.

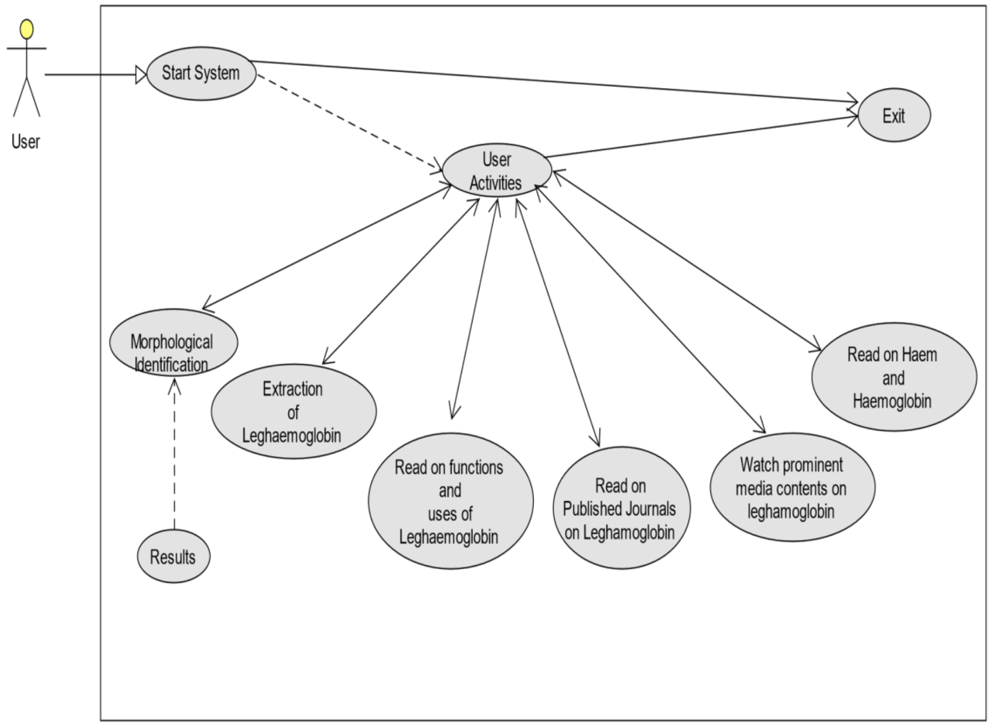

Fig. 4. Use case Diagram of the System

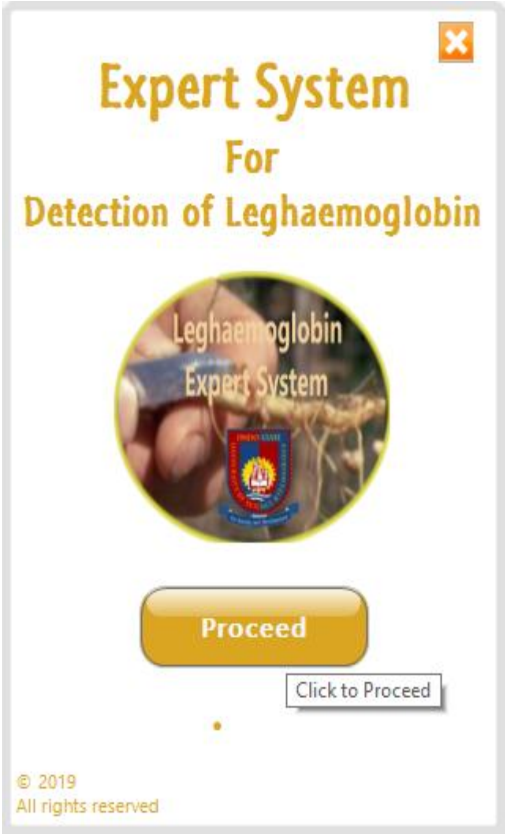

Fig. 5. The system's Welcome Form

Fig. 6. The systems Main-form

Fig. 7. The decision support form

6.5 The Identification of Legume Form

This form makes use of the expert system rule-based inference engine and knowledge base to aid in identifying in hand/ insight plants, to either be leguminous or non-leguminous.

For a user to use the system, he needs to have a sample of the plant in hand or its properties, Fig. 8 depicts how the system uses backward chaining to arrive at conclusions.

This form contains two interfaces which can switch over to either one or the other with a radio button control, the interfaces includes the;

i. Interactive terminal interface: users are asked questions and answers are provided by the users. Inputs by the user are used to make an inference by the inference engine,the same interface is available in the identification of rhizobia form and identification of leghaemoglobin in root nodules form, but with a different knowledge base to infer from.This form is represented in Fig. 9 and Fig. 10.

ii. Selection Panel: selection controls are used on this panel to get input from users to the inference engine to draw an inference and give conclusions. This panel is easier to use by users with little or no knowledge about plants, the same interface is available in the identification of rhizobia form and identification of leghaemoglobin in root nodules form, but with a different knowledge base to infer from. This form is represented in Fig. 11 and Fig. 12.

(Note: only one of this interface can be chosen at a time).

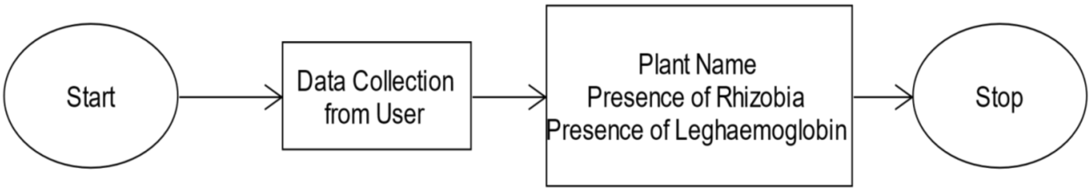

Fig. 8. Backward chain flow

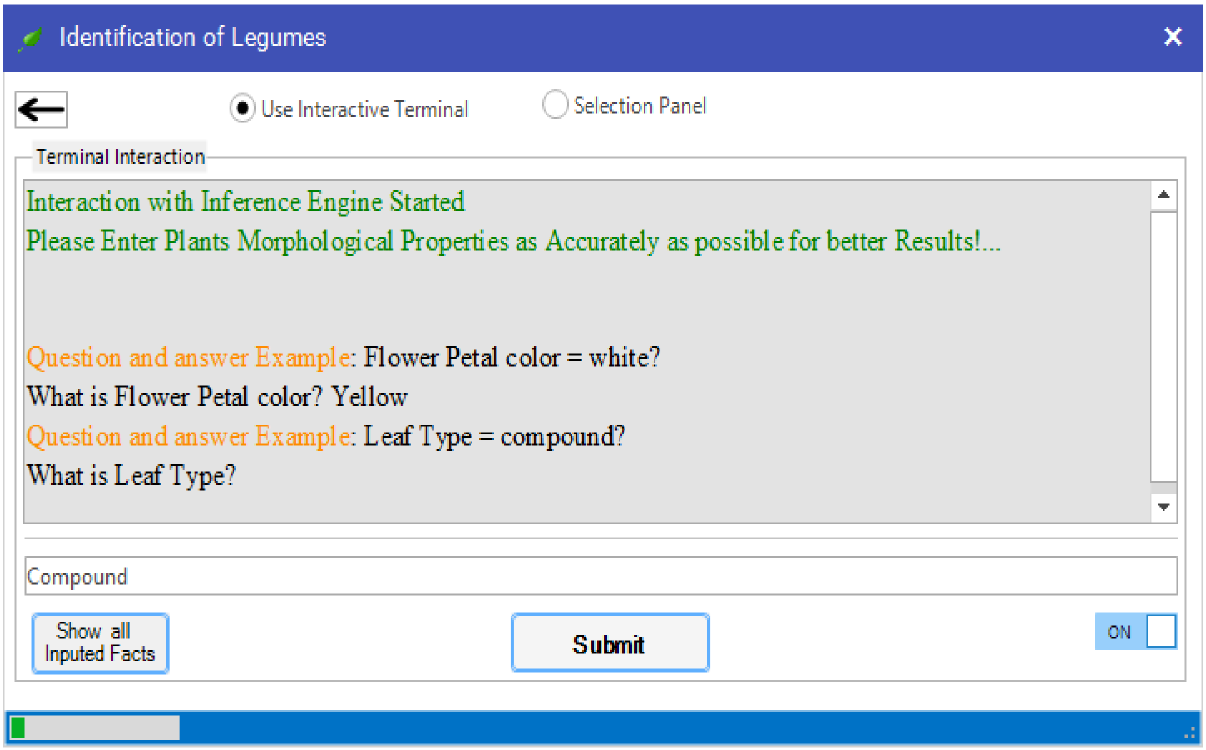

Fig. 9. The identification of legume form interactive terminal interface

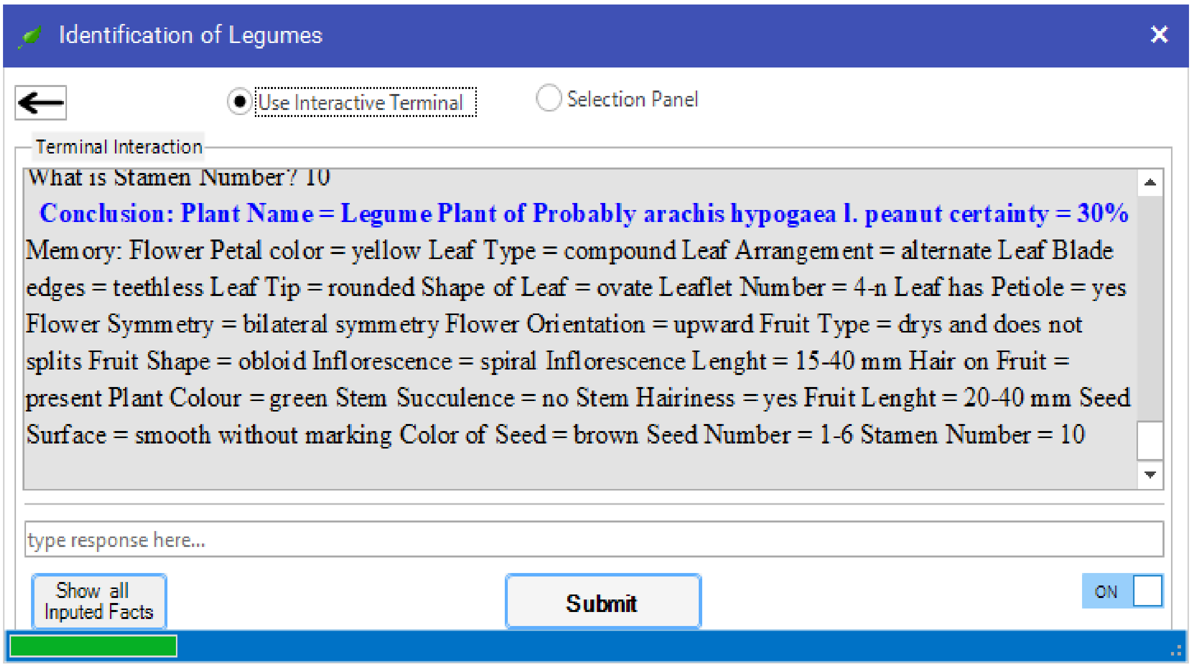

Fig. 10. The identification of legume form interactive terminal interface

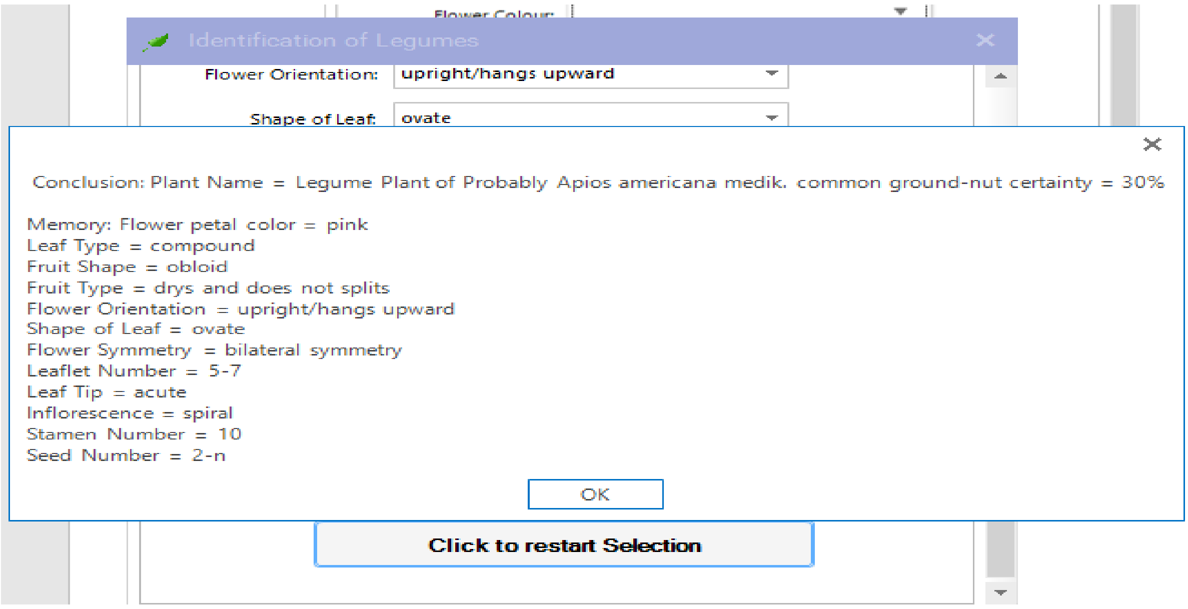

Fig. 11. The identification of legume form Selection panel interface

6.6 The Information Support Panel

This form houses basic information about leghaemoglobin which are important for a user to know and have access to. Information on the extraction of leghaemoglobin molecules is contained in this panel. The form is represented in Fig. 13.

6.7 The Extraction of Leghaemoglobin Form

This form houses information on various steps to be carried out to extract leghaemoglobin from a leguminous plant using soybean as a sample prototype and provides explanations on materials and procedures to be performed in the extraction. This form is represented in Fig. 14.

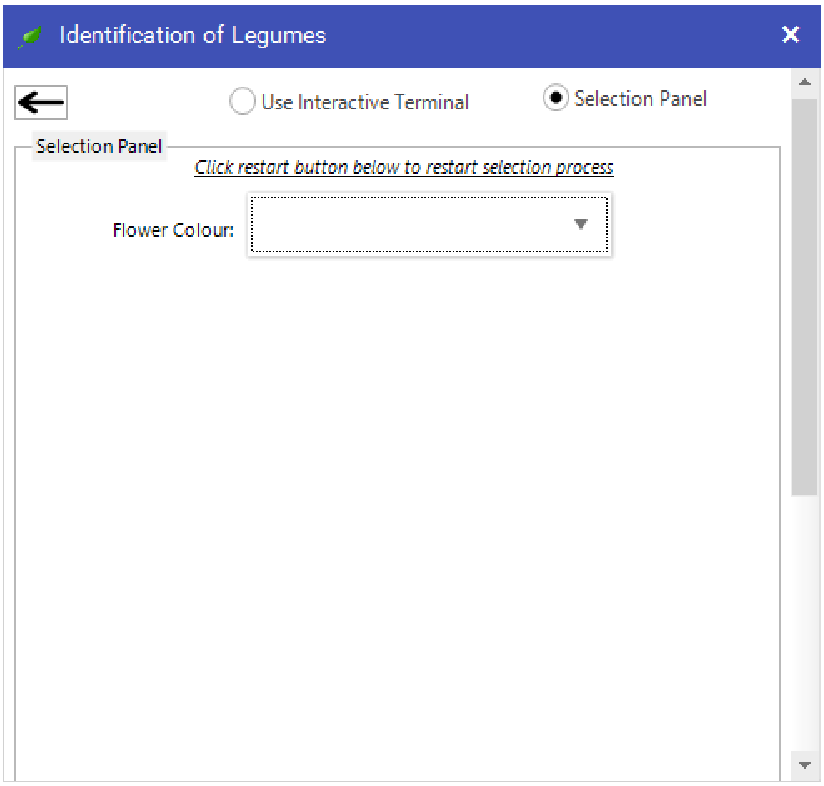

Fig. 12. The identification of legume form Selection panel interface

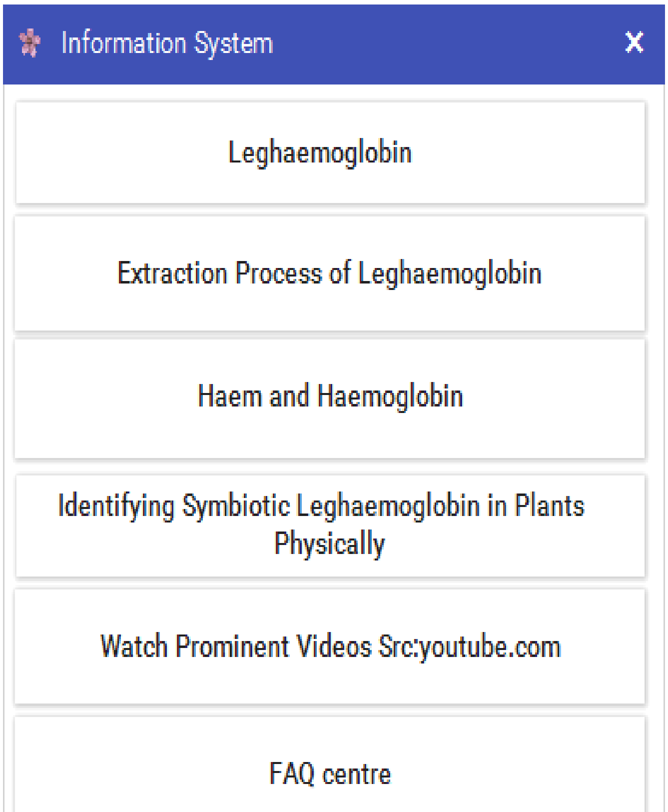

Fig. 13. Information support form

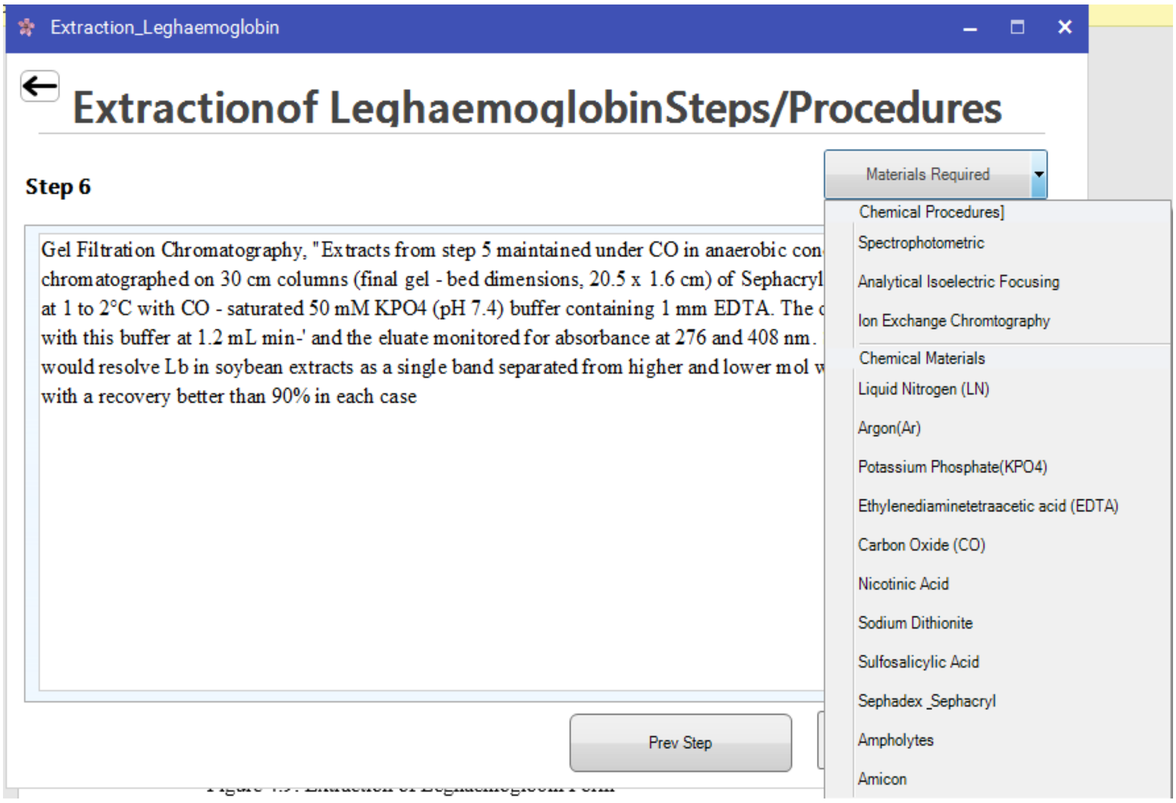

Fig. 14. Extraction of Leghaemoglobin Form

7. CONCLUSION

The developed system is efficient, reliable, cost-effective and has a good response result of an acceptable percentage. It can be utilised by experts and non-experts in getting vital information about leghaemoglobin and aiding as a quick reference in a laboratory setting.

The developed application fulfils all the objectives of the research. It is of great social and ecological importance, given the following:

i. It can help in the identification of legumes, rhizobia and leghaemoglobin morphologically from a PC without the necessity of any network connection or other technology, with a moderate level of accuracy comparable to that of an expert.
ii. It would provide a means of botanical and biochemical knowledge on the subject domain (leghaemoglobin); users would be able to read about and see the various structural diagram of leghaemoglobin. To gain better knowledge about the domain.
iii. It would reduce the instances of misinformation and enhance public knowledge about the leghaemoglobin

It reduces the effort of a human expert to repetitively help in identification, extraction and explanation of the basic concepts of leghaemoglobin, further explanation requires a botanical or biochemical expert

COMPETING INTERESTS

Author has declared that no competing interests exist.

REFERENCES

1. Samy S, Abu N, Abu Z. An expert system for diagnosing eye diseases using CLIPS. Journal of Theoretical and Applied Information Technology. JATIT. 2008;923-930.
2. Mark R. O'brian, Paul M. Kirshbom, Robert J. Maier. Bacterial heme synthesis is required for expression of theleghemoglobinholoprotein but not the apoprotein insoybean root nodules. Proc. Nati. Acad. Sci. USAv. 1987;84:8390-8393.
3. Github.com/chen0040/cs-expert-system-shell. 2017.
   Retrieved February 2018, from Github: Githun.com.
4. Available:Gobotany.newenglandwild.org/search/?q=legumes; 2018.
   Retrieved January 2019,

Available:gobotanynewenglandwild:gobotany.newenglandwild.org
5. Domingo M, Uriz MJ. Design and development of SPONGIA, an Expertsystem forsponges' identification. SCIENTIA MARINASCI. MAR. 1998;62 (1-2):45-57.
6. Amanuel Ayde Ergado. Self learningcomputer troubleshooting expert system. International Journal of Artificial Intelligence & Applications (IJAIA). 2016;7:45-58.
7. Aleksandar Rikalovic, Ilijacosic. A fuzzy expert system for industrial location factor analysis. Acta Polytechnica Hungarica. 2015;12(2).
8. Ayodeji D, Oriyomi A, Ayooluwa, Sunday T. A rule based expert system for diagnosis of fever"; International Journal of Advanced Research. 2013;1(7).
9. Giarratano J, Riley G. Expert systems: Principles and programming. Massachusetts: Thomson Course Technology; 1998.
10. Hatzilygeroudis I, Vassilakos PJ, Tsakalidis. Proceedings of the 1st International Conference on Medical Physics and Biomedical Engineering (MPBE'94), Nicosia, Cyprus. May 1994;I:148-152.
11. Nureize A, Shafiza E, Norlida H, Zehan A. Integrated knowledge based expert system for disease diagnosis system. International Research and Innovation Summit (IRIS2017); 2017.
12. Patra P, Sahu P, Mandal I. An expert system for diagnosis of human diseases. International Journal of Computer Application. 2010; 1(13):71-73.
13. Cuena J, Fernandez G, Lopez de Mantaras R, Verdejo MF. Inteligencia artificial: Sistemas expertos (Alianza Editorial); 1985.
14. Yerokun OM, Onyesolu MO. Developing and evaluating a neuro-fuzzy expert system for improved food and nutrition in Nigeria. Open Access Library Journal. 2021;8:e7315.
DOI:https://doi.org/10.4236/oalib.1107315
15. Jackson Peter. "Introduction to expert system"(3rd Ed.) Addison Wesley. 1999;2.
ISBN 978-0-201-87686-4
16. Garcia N, Perez R, Silverman B. On teaching ai and expert systems courses. IEEE Transactions on Education. 1993;36(1): February.
17. Hustinawaty, Randy Aprianggi. "The Development of Web Based Expert System for Diagnosing Children Diseases Using PHP and MySQL". International Journal of Computer Trends and Technology (IJCTT). 2014;10(4) April 2014.
18. Available:En.wikibooks.org/wiki/Expert_Systems/Components_of_Expert_Systems.
2018. Retrieved January 2019, from wikibooks.com: en.wikibooks.org.