ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

Design and Development of Semantic Web based Guava Expert System for Disease Classification

M. Surendra Prasad Babu[1], Dasari Siva Krishna[2]

Professor, Dept. of CS & SE, Andhra University, Visakhapatnam, India[1]

M.Tech, Dept. of CS & SE, Andhra University, Visakhapatnam India[2]

Abstract: Neural net is an artificial representation of the human brain that tries to simulate the learning process. Artificial Neural Networks (ANN) architecture is a Multilayer Perceptron (MLP) network, widely used for solving problems related to data classifications, implemented in computer programs which involve large number of necessary calculations. Expert Systems are used to simulate human expert for solving a variety of engineering problems. Normally acquiring knowledge is a big problem in developing the expert systems. In this paper a three layered artificial neural network is used for identifying the disease of the Guava crop. A critical study is conducted for understanding the semantic web stack and applied to design and develop a "Semantic Web Based Guava Expert System". The Guava expert system has two modules namely, information system and expert advisor system. The Guava crop Information system provides some semantic information about Guava varieties, nutrient management, pest and disease management, weed management and post harvesting techniques, whereas advisory system takes inputs from the farmer and processes it using back propagation algorithm to identify the disease.

Keywords: *URIs, Resource Description Framework, RDF schema, Ontology, Artificial Neural Networks, Back Propagation Algorithm.*

I. INTRODUCTION

The Semantic web is an extension of the World Wide Web with the new technologies and a set of standards that are recommended by the W3C (World Wide Web Consortium) [1]. The set of standards and best practices for sharing data and the semantics of that data over the web for use by applications. In semantic web, Ontology is consists of definitional aspects such as high-level schemes and assertion aspects (entities, attributes, an interrelationship between entities, domain vocabularies etc.) connect in a semantic manner. Swrl rules, RDF, XML, URI etc., are also implemented in the Semantic Web implementation. It stores the information in the form of tuples. A tuple is represented with the format subject, predicate, object.

In Agricalture field to obtain more value-added products, a product quality control is essential required so that quality of agricultural products increases. Many studies show that there are many causes to reduce the quality of agricultural products reduces. One of the most important factors of such quality is plant diseases. Expert systems are more powerful approach that simulates knowledge of a human expert in certain domain for assisting human to make decisions at a level or greater than human expert. People used to approach experts to seek their advice regarding crop planning which is very expensive and time consuming. If the human expert is not available at that crucial time the farmer may not be in a position to make a right decision for sowing at the schedule time. In this paper, the web based expert system is developed with semantic web technology, which gives better advice thanthe earlier approach. The proposed system uses back propagation algorithm to classify athe diseases based on the SWRL rules generated during the traing process.

II. LITERATURE REVIEW

An expert system is a computer system that emulates the decision-making ability of a human expert. Buchanan, Bruce G., and Edward Hance Shortliffe (2001)[5] discussed Rule-Based expert system for MYCIN. Prof M Surendra Prasad Babu et al. (2010) [6] presented an expert advisor system using an Artificial Bee Colony algorithm for disease management in the Garlic Expert advisory System. They developed the system using JSP as front end and MYSQL as backend. Prof. M. Surendra Prasad Babu et al(2013)[7] developed knowledge base for Soya Bean and expert system using Bagging algorithm to find the disease for the crop. M. S. Prasad Babu et al (2012) [8] designed an expert system for maize crop using Ada-Boost Algorithm and Naive Bayesian Classifier. The system receives the symptoms from the user and processes the information using trained data and determines the diseases of the crop. Heermann, Philip D., and Nahid Khazenie[2] proposed a back-propagation neural network for classification of multispectral image data is explored. This method is a statistical contextual technique, a supervised piecewise linear classifier, and an unsupervised multispectral clustering algorithm. All three techniques were applied to simulated and real satellite imagery. Results from the classification of both Monte Carlo simulation and real imagery are summarized. Prof M S Prasad Babu et al(2011)[10] deals with the design and implementation of an online 'Web Based Exotic Birds Expert Advisory Systems using Back Propagation' Algorithm with Calming Rate. This exotic bird expert advisory system is aimed at identifying the diseases in exotic birds and advises the pet lovers about the proper curing mechanism through online. By using the back propagation algorithm without calming rate about 3000 to

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

4000 iterations be required for finding a disease in an expert system at each time.

III. DESIGN OF SEMANTIC WEB BASED EXPERT SYSTEM

Expert Systems are the computer programs that capture knowledge and allow its dissemination to others. An Expert knowledge base is a repository of knowledge consisting of logical rules and procedures. These knowledge is acquired from a specialist, called domain expert in the area . To implement the expert systems, there is a need of intelligent algorithms to emulates the decision-making ability of a human expert.
A semantic web approach that exhibits a degree of expertise in giving advice to farmers regarding the Guava crop. It provides varieties, Propagation, Nutrient management, soil, harvesting management, etc. The Expert system gives advice to the farmer for a particular disease under certain conditions.

The proposed Expert System consists of the following components:
A. Expert
B. Knowledge Base (OWL file)
C. Inference engine (SWRL rules)
D. Advisory Module
E. User (Growers/Users)
F. User Interface.

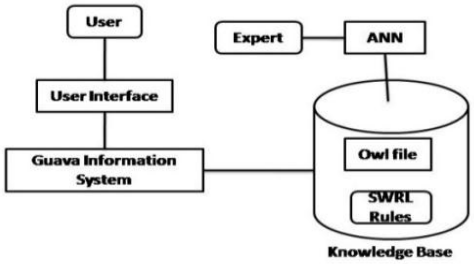

Figure 1. Expert System Architecture for the Guava crop

A. ***Expert:*** Expert is the human who has knowledge and experience about the crop planning for the area. He supplies knowledge to the knowledge base in the form of factual information and in relation to the analytical methods. They use to solve problems in that particular problem field.

B. ***Knowledge Base:*** Knowledge base is the heart of the expert system. The knowledge base contains essential information necessary for understanding, formulating and giving advice for a plantation, and variety to get better yield and better income. The knowledge base is developed by using W3C standards [1] like RDF,RDFS and OWL. The Ontology is modeled using protégé tool. It consists of a set of rules necessary for solving the problem. The knowledge stored in the knowledge base establishes the system capability to act as an expert.

In the proposed system ,semantic knowledge base of the expert system and SWRL rules are used for inference the knowledge base. This information is interpreted by computer in machine readable format. To design the knowledge base W3C (World Wide Web Consortium) proposed semantic web protocol stack, It consists set of standards shown as below.

i. RDF
ii. RDF Schema
iii. Ontology
iv. SWRL Rules

i. ***RDF:*** RDF is a data model, for the representation of resources. A resource can be anything within the context. The web relevance is given to the web resources, that can be located via a URL (Uniform Resource Locator). The URL is one special type of URI (Uniform Resource Identifier). These URL are the basic building blocks of the statements (or tuples). In relational database the data can be represented in terms of tuples (Rows) where as in RDF data can be represented in RDF formats, data flowed in RDF is subject, predicate and Object. Subject and predicates are URIs whereas an object can be a URI or literal. Literal is used to model data values. There are different types in RDF formats, namely N-triples, Turtle, Notation 3, RDF/XML, RDF/JSON, JSON-LD, RDFa etc.

ii. ***RDF Schema:*** RDFS is a Vocabulary description language used in Semantic Web technology, a Vocabulary is a set of terms stored using a standard format that people can reuse. A vocabulary of property names typically has its own namespace to make it easier to use it with another set of data. Vocabularies are usually stored using the RDF schema and OWL standards.

iii. ***Ontology:*** Ontologies are formal definitions of vocabularies [3] that allow defining complex structures as well as new relationships between vocabulary terms. Ontology defines with OWL; Ontologies are used to capture knowledge about some domain of interest. The most recent development in standard ontology languages is OWL from the World Wide Web Consortium (W3C) Like Protégé, OWL makes it possible to describe concepts but it also provides new facilities. It has a richer set of operators - e.g. intersection, union and negation. It is based on a different logical model which makes it possible for concepts to be defined.

C. ***Inference Engine****:* Rule based systems are commonly used in many domains like engineering, medicine,law, internet access authentication, etc. RuleML (Rule Markup) is a markup language for publishing and sharing rule bases on the World Wide Web. RuleML builds a hierarchy of rule sublanguages upon XML, RDF, and OWL. Inference Engine, gives answers to the questions posed by the farmer/user from the knowledge base. In Ontologies SWRL rules are used for inference mechanism.

SWRL Rules:

- SWRL is an acronym for Semantic Web Rule Language.
- SWRL is intended to be the rule language for the Semantic Web.

DOI 10.17148/IJARCCE.2015.49111

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

- SWRL includes a high-level abstract syntax for hornlike rules.
- All rules are expressed in terms of OWL concepts (classes, properties, individuals).

D.***Advisory Module:*** In this module, a neural network based technique, known as the back propagation is implemented for disease classification.The model is tested with the trained weights, based on these weights the expert generates SWRL rules, these rules gives the classification of the disease.

The following steps showed the ANN with back propagation algorithm.

Step 1: Apply the inputs to the network and work out the output – the initial weights were random numbers and initial output could be anything.

Step 2: In output layer ,calculates Error for each node. The following equation is used for calculating the output layer error.

$$\delta_k = O_k(1 - O_k)(O_k - T_k) \quad (1)$$

where

δ_k is the Error at layer k
O_k is the output at layer k
T_k is the targer outpur at layer k

Step 3: In Hidden layer,calculates Error for each node. The following equation is used for calculating the hidden error layer error

$$\delta j = Oj(1 - Oj) \sum_{k \in K} \delta_k W_{jk} \quad (2)$$

where

δ_j is the Error at layer j
O_j is the Expected output at layer j
$W_{jk}$is the weight between layer j and k

Step 4: Update the weights and biases as follows

$$\Delta W = -n\, \delta_l O_{l-1} \quad (3)$$
$$\Delta\theta = n\delta_l \quad (4)$$

Apply the updated weights to the original weights

$$W + \Delta W \rightarrow W \quad (5)$$
$$\theta + \Delta\theta \rightarrow \theta \quad (6)$$

Where

$\Delta W$is the updates weight
$\Delta\theta$ is the updated bais
n is the learning rate
δ_l error at current layer l
O_{l-1} is the previous output.

Step 5: Step 4 repeat until the error is minimized (or) neural network reaches its maximum runs.

E. ***User:*** User is the farmer or any other person who gets the advice to the Guava crop for the disease under certain conditions.

F. ***User Interface:*** The user interface is to exchange information between the user and the inference engine. Web page design techniques are used involving graphical user interface (GUI) controls such as buttons, labels etc.

IV. DEVELOPMENT OF SEMANTIC WEB BASED GUAVA EXPERT SYSTEM

The Guava expert system consists of two modules namely, information system and advisory system. In Information system, Guava crop information is designed using semantic web technology. The information system consists a set of classes and subclasses and each pair of class has a relation to the other classes so that a complete directed graph is designed for the guava crop. In Advisory System, SWRL based inference mechanism and back propagation algorithm is used for disease identification.

A. ***RDF (Resource Description Framework) Model for Guava disease:***

In Guava RDF, The web resources are represented with the URLs. These URLs are the basic buliding blocks for the Guava crop. RDF\XML data model is used for interchange the data.

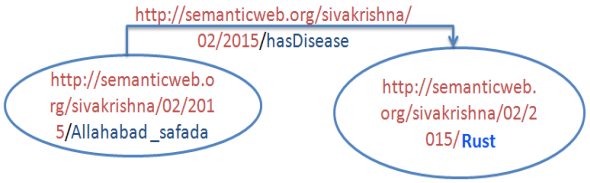

Figure 2. RDF for Guava Disease

The above RDF format shows that the URLs of the subject guava variety, hasdisease is the predicate and the Rust is the Object. The predicate shows the binary relation between the variety of class and the disease variety.

B. ***RDFS (Resource Description Framework Schema) for Guava Disease:***

In Guava RDF Schema, the set of classes like varieties, Diseases, Propagation, Fertilizers etc are the vocabulary terms. These vocabulary terms are used to make relation among the vocabularies.

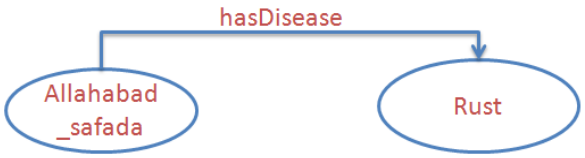

Figure 3. RDF Schema for Guava Disease

The above RDFS shows the vocabulary of the class variety and vocabulary of the Disease. Here Allahabad safada is the individual of the type Guava variety and Rust is the individual of the type disease.

C. ***SWRL Rules (Inference) for Guava Symptoms:***

SWRL rules are used for inference in the knowledge base. a protégé-owl development environment is used for working with SWRL rules.Some of the rules shown below:

Rule1: Part(?p) ∧ hasSymptom(?p, ?sym) → sqwrl:select disease(?p, ?sym)

Rule2:Part(?P)^hasSymptom(?p,?sym) →query:select(?P) ^ query:count(?sym)

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

Rule3: Disease (?p) → query: count (? p)
Rule4: Part(?p) ^hasDisease(?p,?dis)^ hasSymptom(?dis, ?sym) → . select(?P) ^ query:count(?sym).
Rule5:Part(?P)^hasDisease(?p,?sym) →query:select(?P) ^ query:count(?sym)

Description: Rule1 shows that , the variable of the part(p) and the variable of the symptom(sym) it genarates disease Rule2 shows that based the variable of the part(p) and the variable of the symptom(sym) it counts the number of disease.Rule3 shows that, counts number of diseases in disease class. Rule4 shows that,Counts the number of symptoms of the parts.Rule5 shows that, it counts the symptoms of the parts.

D. ***OWL (Web Ontology Language):***

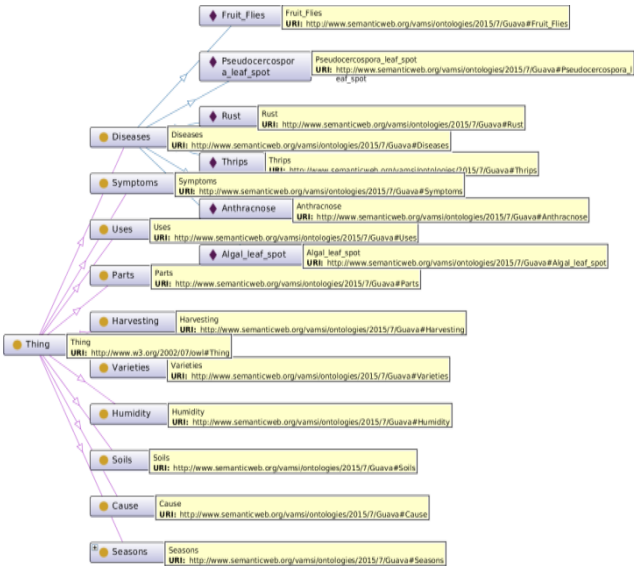

Figure 4. Guava OntoGraph

Guava crop information is designed by using the Web Ontology language. This OWL file is used as a knowledge base for the Expert system. In Guava ontology design, The OWL file consists of one main class and nine sub-classes. Guava is the main class and the sub-class of Guava class are Guava varieties, diseases, Symptoms, Uses, Parts, Harvesting, Soils, Causes, Seasons. For each sub-class there is a list of instances described with the object Property or data Property. Individuals are the instances of the classes.These instances are shown the relation between the classes. The disease class consists of six diseases, each disease is an instance of the diseasse class.

E. Development of ANN for Guava crop:

The Back propagation Algorithm is a learning rule of multi-layered neural networks, credited to Rumelhart and McClelland[10]. The artificial neural network consists of three layers namely, input layer, hidden layer and output layer. Input layer units are connected with the hidden layer as well as hidden layer units are connected to the output layers. The connecting elements called as the weights of the neural network and each unit (neuron) corresponding to the weight passes through the activation function (or) transfer function. In the proposed back propagation algorithm a sigmoid transfer function is used. This function value ranges from (0,1).

$$f(x) = \frac{1}{1+e^{-x}} \qquad (7)$$

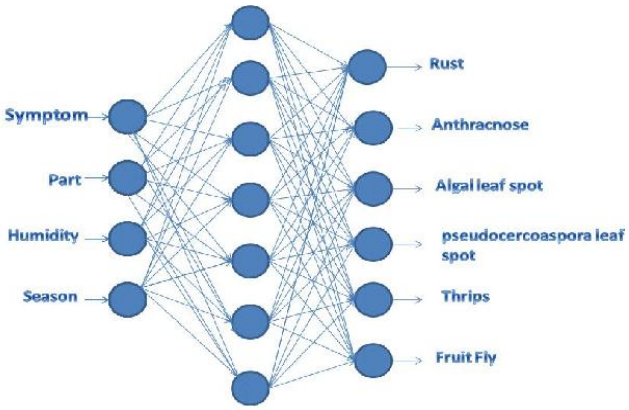

Figure5:Artificial Neural Network

Training and Testing phase:

In this algorithm the neural network consists of sigle input layer with 4 input neurons, 3 hidden layers with 12 neurons in each layer, and single output layer with 6 neurons. The neural network takes these inputs as an input pattern and multiplied with the weights. At the initial stage the weights are random values. These summations pass to the hidden layer neurons. The sigmoid transfer function be used for transfer value from one to another neuron. Thereafter, from hidden layer to output layer also the process will be continued. The algorithm is run until minimize the error=0.001 (or) number of iterations is equal to the 5000 and the threshold($\lambda \geq 0.5$) gives the classification of the guava diseases in the training process. the model is prepared with the training weights and the sample dataset is shown in Table 1

TABLE 1. Sample Dataset

Symptom	Part	Humidity	Season	Disease
Orange Pustules	Leaf	High	Summer	Rust
Dark Colored Lesions	Fruit	High	Summer	Anthracnose
Rust-colored tufts	Leaf	Low	Winter	Algal leaf spot
Dark colored puncture wounds	Fruit	Moderate	Monsoon	Fruit Flies
Orange Pustules	Young Shoots	High	Summer	Rust
Black Wounds	Fruit	Moderate	Autumn	Triples

In testing process the individuals of the OWL file to be tested with input neurons and these values are stored in the OWL literals. The neural network takes these values as an input pattern and fires the one of the output neuron in the output layer.These neurons give the disease

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

classification.The Expert system provides a solution management of the disease to the users/farmers.

V. RESULTS AND DISCUSSIONS

The results of the two modules namely Guava crop information System and Disease Advisory system of the proposed system are shown below.

A.*Report 1*: Information System- Displays the Varieties Of Guava crop.

The information system contains the information regarding Guava varieties, pests, diseases, fertilizers, symptoms, pesticides etc. The users access the requried information about the Guava crop from Information system interface.

Figure 6: Varieties of Guava crop

Discussion: Figure6 displays the list of varieties of Guava, when user select a varieties. They are: Allahabad Safeda,Banarasi,Chittidar,Harijha,RedFleshed,Arka Mridula.

B. *Report 2:* Advisory System- Displays the Disease, cause of the disease and remedial steps to control the disease.

Figure 7a:Test input

Figure 7b: Output Disease and remedial measures

In advisory system(Figure 7a) the user selects the Symptom, part, Humidity range and season of the guava crop. The algorithm process the user request and displays the remedial measures.

C. *Report 3*: Performance analysis of Back Propagation Algorithm

when consider with the 3 hidden layers with 12 neurons of each layer the neural network get an acuracy of 84%. Tesing process 25 test samples are taken for the testing, out of 25 samples 22 are correctly classifified.

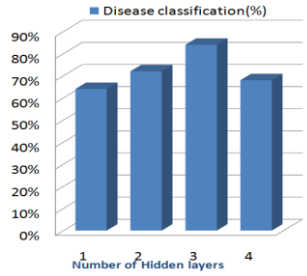

Figure 8: Disease classification

The following result shows that the neural network is tested with the number of hidden layers, when considering with the single hidden layer the algorithm gives 64% disease classification, when considering with the two hidden layers the algorithm gives 72% disease classification, when considering with the three hidden layers the algorithm gives 84% disease classification, when considering with the four hidden layers because of the over fitting the algorithm gives 68% disease classification.

VI. CONCLUSION

A new combination of back propagation algorithm and Semantic web approach are used to develop an expert system for disease classification of Guava crop. This approach facilitates the end user in particular farmers to apply remedial management techniques for obtaining better yield. In the proposed system the Guava Ontology incorporates the knowledge base of the guava expert system and SWRL rules are inference the knowledge base. A translation system and mobile connectivity may be appended for better services to the farmer.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 9, September 2015

REFERENCES

[1]. Tim Berners-Lee, James Hendler and Ora Lassila: "The Semantic Web A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities" Scientific American Feature Article, The Semantic Web, May 2001.
[2]. Heermann, Philip D., and Nahid Khazenie. "Classification of multispectral remote sensing data using a back-propagation neural network." Geoscience and Remote Sensing, IEEE Transactions on 30.1 (1992): 81-88.
[3]. McGuinness, Deborah L., and Frank Van Harmelen. "OWL web ontology language overview." W3C recommendation 10.10 (2004): 2004.
[4]. Myers, Ware. "Introduction to expert systems." IEEE Expert 1.1 (1986): 100-109.
[5]. Buchanan, Bruce G., and Edward Hance Shortliffe, eds. Rule-based expert systems. Vol. 3. Reading, MA: Addison-Wesley, 1984.
[6]. Prof M. Surendra Prasad Babu etal "Implementation of Artificial Bee Colony Algorithm (ABC) On Garlic Expert Advisory System", International Journal of Computer Science and Research, Vol. 1 Issue 1, 2010.
[7]. Prof. M. S. Prasad Babu, Swetha Reddy et al, "A Web-Based Soya Bean Expert System Using Bagging Algorithm with C4.5 Decision trees", International Journal of Agriculture Innovations and Research, Volume 1, Issue 4, ISSN (Online) 2319-1473.Ss
[8]. M.S. Prasad Babu et al, "Development of Maize Expert System Using Ada-Boost Algorithm and Naive Bayesian Classifier", International Journal of Computer Applications Technology and Research, Volume 1-Issue 3, 89-93, 2012
[9]. Mohawk Agarwal, Shivani Gael "Expert System and it's Requirement Engineering Process" proceeding IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014), May 09-11, 2014, Jaipur, India
[10]. K. Srinivasa Rao, and M. Surendra Prasad. "Implementation of Exotic Bird Expert Advisory System using Back Propagation Algorithm with Calming Rate."International Journal of Advanced Research in Computer Science 2.3 (2011).
[11] http://www.indiahorti.net/
[12] http://www.indiakisan.net

.