Medical Expert System on Gradual Loss of Vision (MESGLV) Implemented by Using Induction Rule

Thiri Maung, Swe Zin Hlaing
University of Computer Studies, Yangon
thirimaung84@googlemail.com

Abstract

As the technology grows rapidly, medical diagnosis reasoning is a very important application area of computer-based system. People take a great interest in computer and then computer-based methods are increasingly used to improve the quality of the medical services. Human experts in medical field are frequently in great demand. Thousands of patients are visit to doctors for their loss of vision; these consultations are rewarding as the chief skill is in interpreting the history, not in taking it; as much as allowing it to extend. Therefore, the aim of this paper is to develop a medical expert system to get more quickly decide for the gradual loss of vision. In order to get optimal treatment and care of specialist, the patient medical records must be easily retrieved and the previous data needed for this process must be stored in a certain database beforehand. Induction means a process of reasoning from the specific to the general. In Expert System, it refers to the process in which rules are generated by a computer program from example cases. Induction method use attributes values and selections to rules.

1. Introduction

In medicine, it is very difficult to get correct diagnosis because there are many possible diseases in each case. Moreover, these possible diseases are complicated. And it is also important to get correct diagnosis as soon as possible to save lives of people. Nowadays, medical diagnosis reasoning is a very important application area of computer-based system.

Computer-based methods are increasingly used to improve the quality of medical services. Those methods include both conventional techniques, such as database management systems (DBMSs), and artificial intelligence (AI) techniques. In medical diagnosis, DBMSs are used for storing, retrieving and generally manipulating patient data, whereas expert systems (ESs) are mainly used for performing diagnoses based on patient data.

The main purpose of the paper is to provide an instant online diagnosis of respective symptoms in a user-friendly format, and then predicts the most likely what disease the patient might have and give the advice to the doctor and the system is to aid the doctors in decision making process during the patient's attendance, supplying the most correct diagnosis in relation to the types of respective symptoms.

This paper is organized as follows: Section-2 presents the theory of the expert system, Section-3 presents the proposed system architecture, Section-4 presents the implementation of the system, Section-5 presents the analysis of data and Section-6 presents the conclusion.

2. Background Theory

2.1 Expert System

Expert systems are computerized advisory programs that attempt to imitate the reasoning processes and knowledge of experts in solving specific types of problems.

Thus, an expert system contains a large body of knowledge concerning with one special field that has been provided by one or more human experts in that field and is able to achieve the same performance in problem solving like those experts.

The expert systems are used in two different ways:

1. Decision Support: Reminding information or options to an experienced decision maker. This is commonly used in medicine.
2. Decision making: Allowing an unqualified person to make a decision beyond his or her level or training. This is commonly used in industrial systems.

2.2 Knowledge

Knowledge is basic of the expert systems and it is a collection of specialized facts, procedures, and judgment rules. Knowledge may be collected from many sources. A representative list of sources includes experts in the specific domain, books, films, computer databases, pictures, map, flow diagrams, stories, songs, or observed behavior.

2.3 Knowledge Base

Knowledge engineers, who translate the knowledge of real human experts into rules and strategies, create a knowledge base. Knowledge is represented as a set of rules and data is represented as a set of facts. These rules compare each rule in the knowledge base (rules) with the facts. If a rule processes a fact, the rule is "FIRE", and the "THEN" action (consequence) is executed [2].

2.4 Rule Based

The most common form of architecture used in expert and other types of knowledge-based systems in the production system, also called the rule-based system. This holds the set of rules of inference that are used in system.

Most of these systems use IF-THEN rules to represent knowledge. Rules have an ancestor or condition part, the left-hand, and a conclusion or action part, the right-hand size. These rules are often represented in the following form:

IF
<a set of conditions is true>
THEN
<certain conclusions can be drawn>

2.4.1 Two Types of Rules

Rules are categorized into two types of rule sets, namely ordered rule sets and unordered rule sets.

2.4.1.1 An Ordered Rule

An ordered rule set is referred to as a decision list in which rules are correlated. The following is an example of ordered rule set (Clark and Boswell, 1991),

IF feathers = yes THEN class = bird
ELSE IF legs = two THEN class = human

Ordered rule sets are interpreted in a top-down manner as if-then-else rules. When classifying a new instance, each rule in an ordered rule set is tried in sequential order until one is found whose conditions satisfy the attribute values of the instance. This rule is executed or "fired" and the class of the "fired" rule is assigned to the instance.

2.4.1.2 Unordered Rule

In contrast, an unordered rule set is a set of independent rules. For example, the above ordered rule set can be transformed into an unordered rule set:

IF feathers = yes THEN class = bird
IF feathers = no AND legs = two THEN class = human.

When classifying a new instance with an unordered rule set, all rules that are eligible for firing are collected and the most probable class of all these rules is assigned. Both unordered and ordered rule sets have some problems. In unordered rule sets, rules that satisfy the same conditions may have different class values [4].

2.5 Confirmation Rules

In decision model based on confirmation rules, every diagnosis class is treated separately as the target class. For a given target rule is a conjunction of logical tests (literals). Confirmation rules have a similar form as, for example, induction algorithm and IF-THEN rules [1].

2.6 Induction Rule

Induction means a process of reasoning from the specific to the general. In ES, it refers to the process in which rules are generated by a computer program from example cases. Induction method uses attributes, values and selections to rules. A rule induction system generates decision rules from a set of training data. Data from which rules are usually presented in a form similar to a table in which *cases (or examples) are labels (or names) for rows* and variables are labeled as *attributes and a decision.*

A problem requires making a prediction about an object, such as a piece of text, based on information known about it. Assume that the attributes used to describe the object are *Aj, Ai, ...,* A,„ each of which can take V, values. The product of these attribute values constitutes the instance space containing all possible instances of the object. Typically, after necessary data about an object with an identified target variable (class attribute) and its attributes have been collected.

The evaluation of a rule induction algorithm is usually based on the following criteria:

Accuracy: How well the rule induction algorithm succeeds in reducing error. Accuracy is commonly measured by the percentage of incorrectly classified instances in the data and includes both training and testing error rates.

Computational efficiency: The time needed to learn a classification model, including the time to apply it to new cases [3].

2.7 Induction Rule Algorithm

Input : a training data set D
Output : a rule set R
Algorithm : set R = 0
while (D != 0)
    Find a rule from D
    Remove all data examples covered by the rule from D
    Add the rule to R
Return R

3. Proposed System Architecture

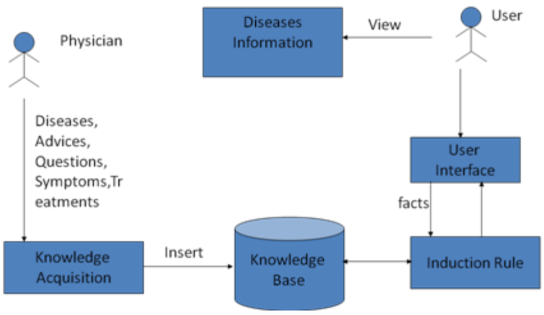

Figure 1. System Design of the proposed system

The above figure shows the architecture of the proposed system. The system stores the patients' information in the form of rule based features in knowledge base. The system can be used by two types of users depending on their roles. For authorized user to maintain the knowledge base, he/she can saved new diseases with their corresponding confirmed solution into the knowledge base. For users, he/she can use the system to support the diagnosis of their input symptoms by entering the disease features. To give the solution of the disease name, the system uses the rule induction algorithm from the knowledge base.

In this phase, the system can retrieve the diagnosis and the treatment to the user by using rule induction algorithm.

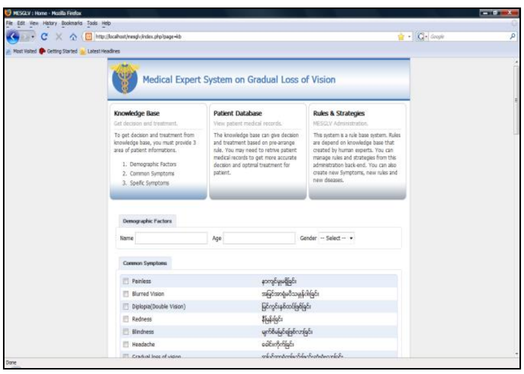

Figure 2. Common Symptoms for the eye diseases

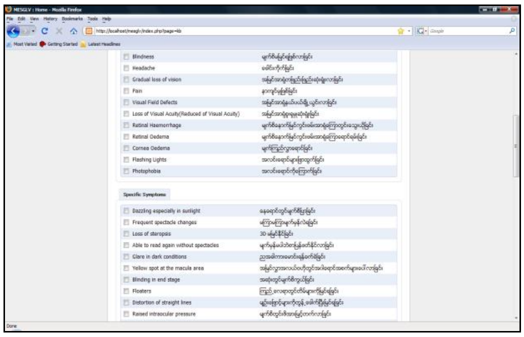

Figure 3. Specific Symptoms for the eye diseases

3.1 Induction Rule Algorithm in diagnosis of Eye Diseases

Rule induction algorithm, already mentioned in Section2.7 is a method of AI, has been used in this diagnostic system. As in all rule based system, the system attempts to produce a solution to diagnosis name by making use of induction algorithm. The diseases that the system used for diagnosis are Cataract, Age related macular degeneration, Glaucoma, Diabetic Retinopathy, Hypertensive retinopathy, Optic atrophy, Slow retinal detachment, Choroidal melanoma, Drug toxicity, Heridatary retinal dystrophies and Sarcoidosis. The system uses the signs and symptoms of each disease as the features of the rules to store in the knowledge base.

4. Implementation of the System

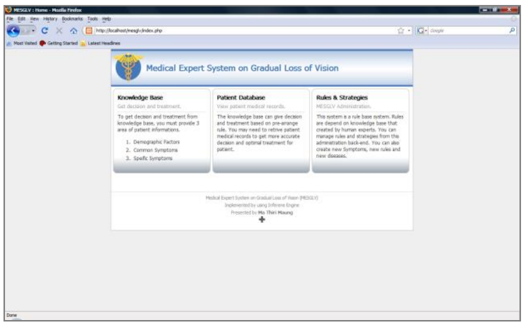

Figure 4 . Main Page of the system

Main page consists of three parts: patient database, rules and strategies for symptoms and knowledge base that stores the symptoms and diagnosis as shown in Figure 4. When the patient enters to the system, the system will be filled the patient's name, age and gender. After filling the required dermographic factors and the symptoms that the patient suffers, it can get "Decision" for the patient by clicking "Get Decision" button as shown in Figure 5.

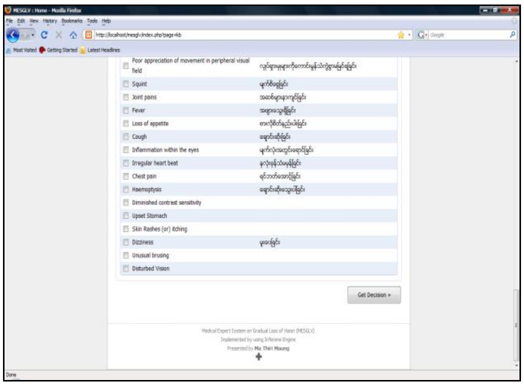

Figure 5.To get decision for the eye diseases

When the system shows the result for the respective symptoms, the result of the disease can be saved in the patient database for further effective treatments in the future as shown in Figure 6.

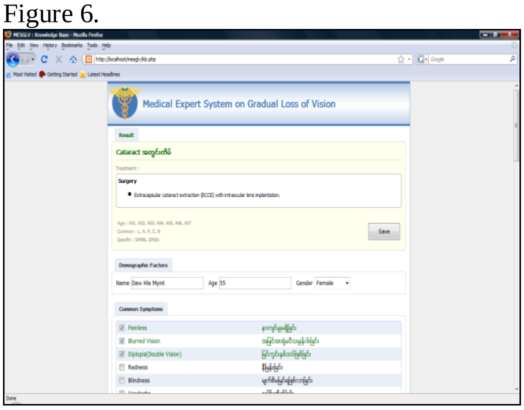

Figure 6.Saving patient's record in the patient database

5. Analysis of Data

The proposed system supports the diagnosis of 11 types of diseases concerning loss of vision namely: Cataract, Age related macular degeneration, Glaucoma, Diabetic Retinopathy, Hypertensive retinopathy, Optic atrophy, Slow retinal detachment, Choroidal melanoma, Drug toxicity, Heridatary retinal dystrophies and Sarcoidosis. The system is developed by 15 common symptoms and 51 specific symptoms of eye diseases for above stated eye diseases. Firstly, a data set is used to train the system. The system is presented with a set of inputs that have known outputs. By comparing output of the system with the known outputs, we can examine the accuracy of the system. The accuracy of the system for diagnosis of each disease is shown in Table 1.

5.1 Classifier Accuracy

Accuracy can be measured by sensitivity, specificity and precision. Sensitivity and specificity are the most widely used statistics used to describe a diagnostic test. Sensitivity is the probability of a positive test among patients with disease. Specificity is the probability of a negative test among patients without disease.

$$\text{Sensitivity} = \frac{t\_pos}{pos} = \frac{\text{Number of true positive}}{\text{Number of positive}}$$

$$\text{Specificity} = \frac{t\_neg}{neg} = \frac{\text{Number of true negative}}{\text{Number of negative}}$$

$$\text{Accuracy} = \text{Sensitivity}\frac{pos}{pos+neg} + \text{Specificity}\frac{neg}{pos+neg}$$

where,
t_pos= no of true positives
pos = no of positive
t_neg= no of true negatives
neg = no of negative
f_pos = no of false positives

Table 1. The accuracy for each disease

	Diseases	Accuracy
1	Cataract	80%
2	Age related macular degeneration	80%
3	Glaucoma	93.33%
4	Diabetic retinopathy	66.67%
5	Hypertensive retinopathy	80%
6	Optic atrophy	73.33%
7	Slow retinal detachment	80%
8	Choroidal melanoma	66.67%
9	Drug toxicity	80%
10	Heridatary retinal dystrophies	73.33%
11	Sarcoidosis	66.67%

Among these all diseases, Glaucoma, Cataract, Age related macular degeneration and Drug toxicity are possibility of diseases and more frequently occur In Myanmar than the others.

6. Conclusion

By using this system, the physicians will check easily the patients' detailed medical information and can easily know how to treatment this patient. This thesis is focused on the processing of induction rule, knowledge base and expert system. Furthermore, the physicians can reduce time consuming to find the diagnosis for the patients symptoms by reviewing previous information and can give treatments quickly to the patients. From the perspective of the patients, they can review the past diagnosis and can get treatments, advices and more consultations quickly from the physicians.

7. References

[1] G.Dragan, L.Nada, J.Viktor, "High Confidence Association rules for Medical Diagnosis", Journal of Computing, 2002.

[2] T. Efraimm, "Expert Systems and Applied Artificial Intelligence", ISBN 0-02-946565-6

[3] S. Rainer, V. Olga, "Adaption and Case-Based Reasoning Focusing on Endocrine Therapy Support", Connecting Medical Informatics and Bio-Informatics, ENMI, 2005

[4] Zaw, Wint Thida, "Diagnosis of the Diseases on Respiratory Symptoms Using the Expert System Approach", 2003

[5] X. Xuepene, "Classification Rule Induction with an Ant Colony Optimization Algorithm", Journal of Computing, August 2004