Chapter 3

Neural Networks and Decision Trees For Eye Diseases Diagnosis

L. G. Kabari and E. O. Nwachukwu

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/51380

1. Introduction

Clinical Decision Support Systems (CDSS) provide clinicians, staff, patients, and other individuals with knowledge and person-specific information, intelligently filtered and presented at appropriate times, to enhance health and health care [1]. Medical errors have already become the universal matter of international society. In 1999, IOM (American Institute of Medicine) published a report "To err is Human" [2], that indicated: First, the quantity of medical errors is incredible, the medical errors had already became the fifth lethal; Second, the most of medical errors occurred by the human factor which could be avoid via the computer system. Improving the quality of healthcare, reducing medical errors, and guarantying the safety of patients are the most serious duty of the hospital. The clinical guideline can enhance the security and quality of clinical diagnosis and treatment, its importance already obtained widespread approval [3]. In 1990, clinical practice guidelines were defined as "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances" [4].

The clinical decision support system (CDSS) is any piece of software that takes as input information about a clinical situation and that produces as output inferences that can assist practitioners in their decision making and that would be judged as "intelligent" by the program's users [5].

Artificial intelligence has been successfully applied in medical diagnosis. They have been used for skin disease diagnosis, fetal delivery, metabolic synthesis as demonstrated in [6,7 and 8]. Artificial neural networks are artificial intelligence paradigms; they are machine learning tools which are loosely modelled after biological neural systems. They learn by training from past experience data and make generalization on unseen data. They have been applied as tools for modelling and problem solving in real world applications such as speech recognition, gesture recognition, financial prediction, and medical diagnostics [9, 10, 11 and 12]. Backpropagation employs gradient descent learning and is the most popular algorithm used for training neural networks. Neural networks were recently viewed as 'black boxes' as they could not explain how they arrived to a particular solution. Knowledge extraction is the process of extracting valuable information from trained neural networks in the form of 'if-then' rules as shown in [13 and 14]. The extracted rules describe the knowledge acquired by neural networks while learning from examples.

The human eye is the organ which gives us the sense of sight allowing us to learn more about the surrounding world than we do with any of the other four senses. We use our eyes in almost every activity we perform whether reading, working, watching television, writing a letter, driving a car and in countless other ways. Most people probably would agree that sight is the sense they value more than all the rest.

A recent survey of 1,000 adults shows that nearly half - 47% - worry more about losing their sight than about losing their memory and their ability to walk or hear. But almost 30% indicated that they don't get their eyes checked. Many Americans are unaware of the warning signs of eye diseases and conditions that could cause damage and blindness if not detected and treated soon enough.

In spite of the high prevalence of vision disorders in this country, so far, few victims receive professional eye care due to one of the following reasons;

- Specialist in eye diseases(ophthalmologist) are few and ophthalmology clinic are also few
- Lack of knowledge that early professional eye care is needed when symptoms are suspected.
- Inability to pay for the needed services.

Due to all of these, late detection of vision disorders and unnecessary loss of vision is encountered. But with a computer based system (expert system), over dependence on human expert can be minimized. This will go a long way to save time and furthermore early detection of eye disease can be adequately addressed. Cost for the services can also be reduced as a lot of unnecessary laboratory test may be avoided with the use of the proposed system.

This study classifies eye diseases using patient complaint, symptoms and physical eye examinations. The disease covered includes the following eye disease; Pink eye (conjunctivitis), Uveitis, Glaucoma, Cataract, Macular Degeneration, retinal detachment, Corneal ulcer, Color blindness, Far sightedness(hyperopia), Near sighteness(myopia), and Astigmatism.

We train artificial neural networks to classify eye diseases according to patient complain, symptoms and physical eye examination. We then use decision trees to extract knowledge from trained neural networks in order to understand the knowledge represented by the trained networks. Finally, we apply decision trees to build a tree structure for classification on the same sets of data sample we used to train neural networks earlier. In this way we combine neural networks and decision trees through training and knowledge extraction. The extracted knowledge from neural networks is transformed as rules which will help ex-

http://dx.doi.org/10.5772/51380

perts in understanding which combination of symptom, physical eye examination and patient's complain constituents have a major role in the eye problem. The rules contain information for sorting eye diseases according to their symptoms, physical condition and complain from the patient and knowledge acquired by neural networks from training on previous samples.

2. Application of Neural network in Clinical decision Support System

These days the Artificial Neural Networks(ANN) have been widely used as tools for solving many decisions modeling problems. The various capabilities and properties of ANN like Non-parametric, Non-linearity, Input-Output mapping, Adaptivity make it a better alternative for solving massively parallel distributive structure and complex task in comparison of statistical techniques, where rigid assumptions are made about the model. Artificial Neural Networks being non-parametric, makes no assumption about the distribution of data and thus capable of "letting the data speak for itself". As a result, they are natural choice for modeling complex medical problems where large database of relevant medical information are available [15].

In biomedicine, the assessment of vital functions of the body often requires noninvasive measurements, processing and analysis of physiological signals. Examples of physiological signals found in biomedicine include the electrical activity of the brain-the electroencephalogram (EEG), the electrical activity of the heart-the electrocardiogram (ECG), the electrical activity of the eye-i.e. PERG and EOG-respiratory signals, blood pressure and temperature signals [16].

Often, biomedical data are not well behaved. They vary from person to person, and are affected by factors such as medication, environmental conditions, age, weight, mental and physical state. Consequently, clinical expertise is often required for a proper analysis and interpretation of medical data. This has led to the integration of signal processing with intelligent techniques such as artificial neural networks (ANN), expert systems and fuzzy logic to improve performance [16].

ANN has been proposed as a reasoning tool to support clinical decision-making since 1959 [17]. Some problems encountered have led to significant developments in computer science, but it was only during the last decade of the last century that decision support systems have been routinely used in clinical practice on a significant scale [16].

The literature reports several applications of ANNs to the recognition of a particular pathology. For example, cancer diagnosis [18 and 19], automatic recognition of alertness and drowsiness from electroencephalography [20], predictions of coronary artery stenosis [21], analysis of Doppler shift signals [22 and 23], classify and predict the progression of thyroid-associated ophthalmopathy [24], diabetic retinopathy classification [25], saccade detection in EOG recordings [26] and PERG classification [22].

In this research we apply a hybrid of Neural Network and decision Tree to classify eye diseases according to patient complain, symptoms and physical eye examination. The aim is to help the ophthalmologist interpret the output of the examination systems easily and diagnose the problem accurately [27-29].

2.1. Artificial Neural Networks

Artificial Neural networks learn by training on past experience using an algorithm which modifies the interconnection weight links as directed by a learning objective for a particular application. A *neuron* is a single processing unit which computes the weighted sum of its inputs. The output of the network relies on cooperation of the individual neurons. The learnt knowledge is distributed over the trained networks weights. Neural networks are characterized into feedforward and recurrent neural networks. Neural networks are capable of performing tasks that include pattern classification, function approximation, prediction or forecasting, clustering or categorization, time series prediction, optimization, and control. Feedforward networks contain an input layer, one or many hidden layers and an output layer. Fig. 1 shows the architecture of a feedforward network. Equation (1) shows the dynamics of a feedforward network.

$$S^{l}{}_{j} = g_i\left(\sum_{i=1}^{m} S_i^{l-1} W_{ji}^{l} - \theta_j^{l}\right) \tag{1}$$

where $S^{l}{}_{j}$ is the output of the neuron j in layer$l$, $S_i^{l-1}$is the output of neuron j in layer l - 1 (containing m neurons) and W_{ji}^{l} the weight associated with that connection with j. $\theta_j^{l}$is the internal threshold/bias of the neuron and g_i is the sigmoidal discriminant function.

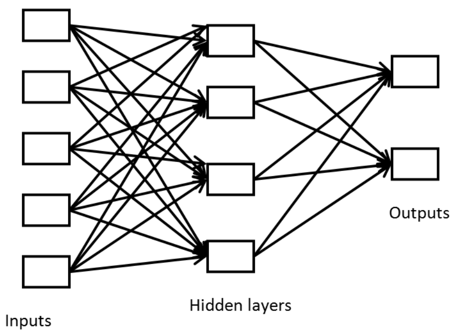

Figure 1. The architecture of the feedforward neural network with one hidden layer.

Backpropagation is the most widely applied learning algorithm for neural networks. It learns the weights for a multilayer network, given a network with a fixed set of weights and interconnections. Backpropagation employs gradient descent to minimize the squared error

http://dx.doi.org/10.5772/51380

between the networks *output values* and *desired values* for those outputs. The goal of gradient descent learning is to minimize the sum of squared errors by propagating error signals backward through the network architecture upon the presentation of training samples from the training set. These error signals are used to calculate the *weight* updates which represent the knowledge learnt in the network. The performance of backpropagation can be improved by adding a momentum term and training multiple networks with the same data but different small random initializations prior to training. In gradient descent search for a solution, the network searches through a weight space of errors. A limitation of gradient descent is that it may get trapped in a local minimum easily. This may prove costly in terms for network training and generalization performance.

In the past, research has been done to improve the training performance of neural networks which has significance on its generalization. Symbolic or expert knowledge is inserted into neural networks prior to training for better training and generalization performance as demonstrated in [13]. The generalization ability of neural networks is an important measure of its performance as it indicates the accuracy of the trained network when presented with data not present in the training set. A poor choice of the network architecture i.e. the number of neurons in the hidden layer will result in poor generalization even with optimal values of its weights after training. Until recently neural networks were viewed as black boxes because they could not explain the knowledge learnt in the training process. The extraction of rules from neural networks shows how they arrived to a particular solution after training.

2.2. Knowledge Extraction from Neural Networks: Combining Neural Networks with Decision trees

In applications like credit approval and medical diagnosis, explaining the reasoning of the neural network is important. The major criticism against neural networks in such domains is that the decision making process of neural networks is difficult to understand. This is because the knowledge in the neural network is stored as real-valued parameters (weights and biases) of the network, the knowledge is encoded in distributed fashion and the mapping learnt by the network could be non-linear as well as non-monotonic. One may wonder why neural networks should be used when comprehensibility is an important issue. The reason is that predictive accuracy is also very important and neural networks have an appropriate inductive bias for many machine learning domains. The predictive accuracies obtained with neural networks are often significantly higher than those obtained with other learning paradigms, particularly decision trees.

Decision trees have been preferred when a good understanding of the decision process is essential such as in medical diagnosis. Decision tree algorithms execute fast, are able to handle a high number of records with a high number of fields with predictable response times, handle both symbolic and numerical data well and are better understood and can easily be translated into if-then-else rules.

The goal of knowledge extraction is to find the knowledge stored in the network's weights in symbolic form. One main concern is the fidelity of the extraction process, i.e. how accurately the extracted knowledge corresponds to the knowledge stored in the network. There are two main approaches for knowledge extraction from trained neural networks: (1) extraction of 'if-then' rules by clustering the activation values of hidden state neurons and (2) the application of machine learning methods such as decision trees on the observation of input-output mappings of the trained network when presented with data. We will use decision trees for the extraction of rules from trained neural networks. The extracted rules will explain the classification and categorization of different eye diseases according to symptoms.

In knowledge extraction using decision trees, the network is initially trained with the training data set. After successful training and testing, the network is presented with another data set which only contains inputs samples. Then the generalisation made by the network upon the presentation is noted with each corresponding input sample in this data set. In this way, we are able to obtain a data set with input-output mappings made by the trained network. The generalisation made by the output of the network is an indirect measure of the knowledge acquired by the network in the training process. Finally, the decision tree algorithm is applied to the input-output mappings to extract rules in the form of trees.

Decision trees are machine learning tools for building a tree structure from a training dataset of instances which can predict a classification given unseen instances. A decision tree learns by starting at the root node and selects the best attributes which splits the training data. The root node then grows unique child nodes using an entropy function to measure the information gained from the training data. This process continues until the tree structure is able to describe the given data set. Compared to neural networks, they can explain how they arrive to a particular solution. We will use decision trees to extract rules from the trained neural networks.

2.3. Decision Tree

A decision tree(DT) is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm. Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal. Another use of decision trees is as a descriptive means for calculating conditional probabilities.

Decision tree learning is a method commonly used in data mining. The goal is to create a model that predicts the value of a target variable based on several input variables. Each interior node corresponds to one of the input variables; there are edges to children for each of the possible values of that input variable. Each leaf represents a value of the target variable given the values of the input variables represented by the path from the root to the leaf.

A tree can be "learned" by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called recursive partitioning. The recursion is completed when the subset at a node all has the same value of the target variable, or when splitting no longer adds value to the predictions. In data mining, trees can be described also as the combination of mathematical and computational techniques to aid the description, categorisation and generalisation of a given set of data. Data comes in records of the form:

$$(x, Y) = (x_1, x_2, x_3, \ldots, x_k, Y) \tag{2}$$

The dependent variable, Y, is the target variable that we are trying to understand, classify or generalise. The vector **x** is composed of the input variables, x_1, x_2, x_3 etc., that are used for that task.

DT offers a structured way of decision making [29,30]. A DT is characterized by an ordered set of nodes. Each of the internal nodes is associated with a decision function of one or more features.. DT approach can generate *if -then* rules. Specific DT methods include Classification and Regression Trees (CART), Chi Square Automatic Interaction Detection (CHAID), ID3 and C4.5. C4.5 which is the extension of ID3[31,32] is very useful in this work. C4.5 Decision Tree is based on Information theory, that is it uses information theory to select features which give the greatest information gain or decrease of entropy [31]. Information gain is the informational value of creating a branch in a decision tree based on the given attribute using entropy theory.

2.4. Anatomy of the Eye

The eye is made up of numerous components. Figure 1 shows the anatomy of the eye.

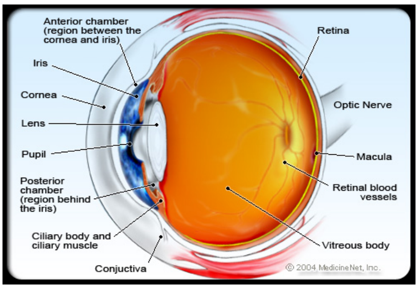

Figure 2. Anatomy of the eye (Source: http://www.medicinenet.com/eye_diseases_pictur_slideshow/article.htm#).

- Cornea: clear front window of the eye that transmits and focuses light into the eye
- Iris: colored part of the eye that helps regulate the amount of light that enters
- Pupil: dark aperture in the iris that allows light to go through into the back of the eye
- Lens: transparent structure inside the eye that focuses light rays onto the retina
- Retina: nerve layer that lines the back of the eye, senses light, undergoes complex chemical changes, and creates electrical impulses that travel through the optic nerve to the brain

- Macula: small central area in the retina that contains special light-sensitive cells and allows us to see fine details clearly
- Optic nerve: connects the eye to the brain and carries the electrical impulses formed by the retina to the visual cortex of the brain
- Vitreous: clear, jelly-like substance that fills the middle of the eye

2.5. Some eye disease conditions

Some eye disease conditions are shown in the figure3 to figure10 below:

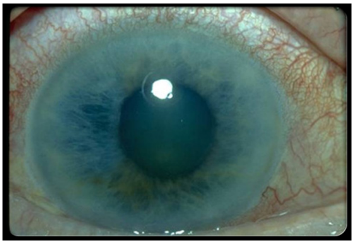

Figure 3. Glaucoma (Source: http://www.medicinenet.com/eye_diseases_pictur_slideshow/article.htm#).

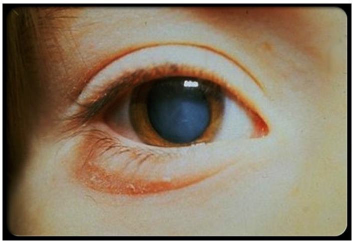

Figure 4. Cataract (Source: http://www.medicinenet.com/eye_diseases_pictur_slideshow/article.htm#).

Figure 5. Macular degeneration (Source: http://www.medicinenet.com/eye_diseases_pictur_slideshow/article.htm#).

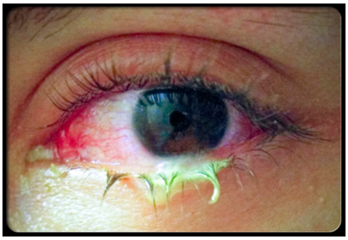

Figure 6. Conjunctivitis (Source: http://www.medicinenet.com/eye_diseases_pictur_slideshow/article.htm#).

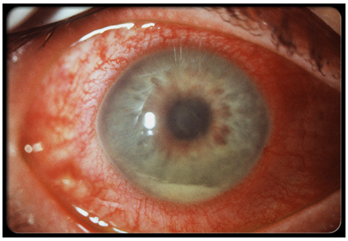

Figure 7. Uveitis (Source: http://www.medicinenet.com/eye_diseases_pictur_slideshow/article.htm#).

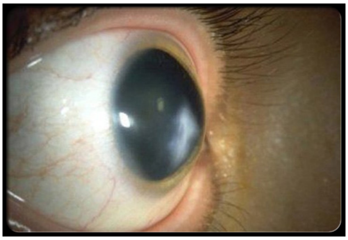

Figure 8. Keratoconus (Source: http://www.medicinenet.com/eye_diseases_pictur_slideshow/article.htm#).

Figure 9. Blepharitis (Source: http://www.medicinenet.com/eye_diseases_pictur_slideshow/article.htm#).

Figure 10. Corneal ulcer (Source: http://www.medicinenet.com/eye_diseases_pictur_slideshow/article.htm#).

3. Eye Diseases Diagnosis

We developed a clinical decision support system which bases its diagnosis on the patient complain, symptoms and physical eye examinations, and uses multilayer feedforward networks with a single hidden layer. Backpropagation algorithm is employed for training the networks in a supervised mode.

The eye diseases selected for diagnosis are as shown in table1. The designed neural network consists mainly of three layers: an input layer, a single hidden layer, and an output layer. The input layer has a total of 22 inputs plus the fixed bias input. These inputs consist of patient complaint, symptoms and physical eye examinations as may either be observed by the ophthalmologist or complained by the patients (i.e. X1, X2,..., X22). The output layer consists of 12 outputs indicating the diagnosed diseases (i.e. d1, d2,..., d13). Table1 shows the selected eye diseases for diagnosis and their symptom and signs as may be complained by patient or observed by the specialist while table2 shows the input variables for the system.

We ran 10 trial experiments with randomly selected 80% of the available data for training and the remaining 20% for testing the networks generalization performance. The learning rate of the network in gradient descent learning was 0.5. The network topology used was as follows: 22 neurons in the input layer for each symptom and signs for the eye disease, 9 neurons in the hidden layer and 13 neurons in the output layer representing each eye disease as shown in Fig. 2. We carried out some sample experiments on the number of hidden neurons to be used in the networks for this application. The results demonstrate that 9 neurons in the hidden layer were sufficient for the network to learn the training samples. The neural network was trained until one of the three following stopping criteria was satisfied:

1. On 100% of the training examples, the activation of every output unit was within 0.2 of the desired output, or
2. a network had been trained for 500 epochs, or
3. a network classified at least 98% of the training examples correctly, but had not improved it's ability to classify the training after ten epochs

http://dx.doi.org/10.5772/51380

S/N	Eye Disease	Signs (patient complain, symptoms and eye condition or cause)
1	Cataracts	Loss of visual acuity, loss of contrast sensitivity, contours shadows and color vision are less vivid, advanced age, bright light or antiglare sunglass may improve vision, poor night vision.
2	Glaucoma	Painless, decrease in peripheral field of view, halos around light, redness of eye, hereditary, aging may also cause it.
3	Macular Degeneration	Blurred vision, distorted images, missing letters in words, difficulty in reading, Trouble discerning colors, slow recovery of visual function after exposure to bright light, loss in contrast sensitivity, advanced age(66-74), Hereditary.
4	Pink eye(conjunctivitis)	Red or pink color eye, itching, blurred image, gritty feeling, irritation, watering of eye
5	Uveitis	Redness of eye, blurred vision, sensitivity to light(photophobia), dark floating sport in visual field, eye pain, blurred vision improves with blinking, discomfort after long period of concentrated use of eye(watching television, using computer or reading).
6	Retinal detachment	Experience of flashes of light and floater in visual view, feeling heaviness in the eye, central visual loss, blind spot in view.
7	Corneal ulcer	Redness of eye, pains of foreign bodies in the eye, pus/thick discharge from the eye, blurred vision, sensitivity to bright light, swollen eyelid, white or grey round spot on the cornea.
8	Keratoconus	Distorted vision, loss of vision focus, contact less could not improve vision
9	Blepharitis	Burning of foreign bodies sensation, itching, sensitivity to light, redness of eye, red and swollen eyelid, blurred vision, dry eye.
10	Color blindness	Problem discerning colors, hereditary, aging.
11	Farsightedness(hyperopia)	Blurred vision for close object, aging, contact lens may improves vision
12	Nearsightedness(myopia)	Blurred vision at distant, good vision for close object.
13	Astigmatism	Blurred vision, steamy appearing cornea, hereditary, may be corrected with contact lens

Table 1. Eye Diseases and their signs (patient complain, symptoms and eye condition).

Input variables	Variable Meaning
X1	Pains in the eye
X2	Redness or pink color of eye
X3	Bright light or antiglare sunglasses improves vision
X4	Poor night vision
X5	Family histories of the eye problem
X6	Decrease in peripheral field of view
X7	Age greater than 45 years
X8	Blurred vision
X9	Blurred vision improves with eye blinking
X10	Distorted vision
X11	Cloudy substance formed in front of eye lens
X12	Slow recovery of vision after exposure to bright light
X13	Irritation, itchy, scratchy or burning sensation of eye
X14	Discomfort after long concentration use of eye
X15	Trouble discerning colors
X16	Floaters in eye, flashes of light, halos around light
X17	Watering or discharge from eye
X18	Swelling of eye
X19	Steamy appearing cornea of eye
X20	Sensitivity to light (photophobia)
X21	Blurred vision for distant objects
X22	Blurred vision for close objects

Table 2. Input Variables and their Meaning.

The backpropagation algorithm with supervised learning was used, which means that we provide the algorithm with examples of the inputs and outputs we want the network to compute, and then the error (difference between actual and expected results) is calculated. The idea is to reduce this error, until the ANN learns the training data. The training begins with random weights, and the goal is to adjust them so that the error will be minimal. The activation function of the artificial neurons in ANNs implementing the backpropagation algorithm is given as follows[33]:

$$A_j(\bar{x}, \bar{w}) = \sum_{i=0}^{n} x_i \cdot w_{ji} \tag{3}$$

$$O_j(\bar{x},\ \bar{w}) = \frac{1}{\left[1 + e^{A_j}(\bar{x},\ \bar{w})\right]} \tag{4}$$

$$E_j(\bar{x},\ \bar{w},\ d) = \sum \left(O_j(\bar{x},\ \bar{w}) - d_j\right)^2 \tag{5}$$

$$\Delta w_{ji} = \eta\left(\frac{\partial E}{\partial w_{ji}}\right) \tag{6}$$

Where: $x_i$are the inputs, $w_{ji}$are the weights, $O_j(x,\ w)$are the actual outputs, $d_j$are the expected outputs and η - learning rate.

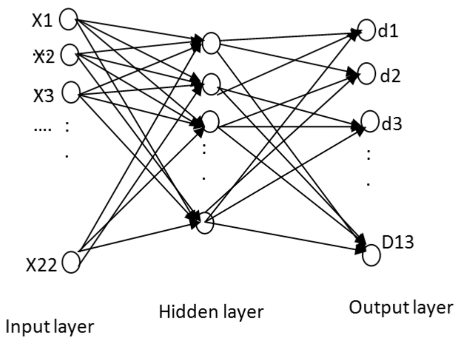

Figure 11. The neural network topology used for diagnosing the eye diseases which contain attribute information of 22 signs and symptoms. Each neuron in the input layer represents a particular sign or symptom. The neurons in the output layer represent the eye disease. Please note that all weight link interconnections are not shown in this diagram.

In this work, we used the C++ programming language in programming neural networks. Data mining and machine learning software tools such as 'Weka' can also be used for classification using neural networks.

4. Experimental Results

4.1. Data Set

The data set used for the training and testing of the system was collected from Linsolar Eye Clinic, Port Harcourt and Odadiki eye clinic, Port Harcourt all in Nigeria. The total data is 400 from which 320 samples (80%) are randomly chosen and used as training patterns and tested with 80 instances (20%) of the same data set. The data set consist of evenly distributed men and women. Samples also consider age randomly collected from 18 years to 70 years.

4.2. Rule Extraction from the ANN

As in Figure 12 it can be seen that both decision trees and neural networks can be easily converted into *IF THEN Rules* or we can simply convert neural networks into decision trees. In this work we use the networks architecture as shown in figure11 together with backpropagation algorithm with supervise learning.

Decision trees are machine learning tool for building a tree structure from a training dataset. A Decision tree learns by starting at the root node and select the best attributes which splits the training data [13]. Compared to neural networks they can explain how they arrive to a particular solution [34]. Hence, it usefulness in clinical decision support system as it may be use to support the expert in his delicate decision making or use as training tools for younger ophthalmologists. A typical decision tree extracted from the neural network in this work is shown in Figure 13.

To simplify complicated drawing the input variables that was shown in table1 may be combined to form conjunctions and negations which may also be used to generate the Decision Tree for some of the eye diseases as shown in Table 3.

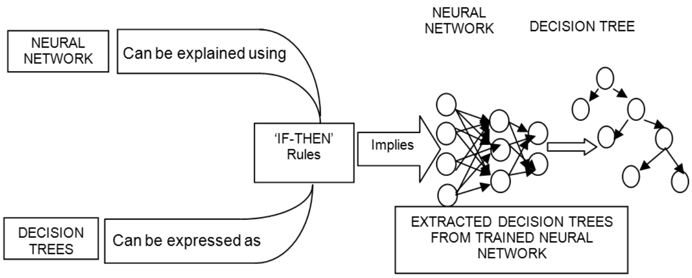

Figure 12. Extracting decision trees from neural networks.

Input variables	Variable Meaning	Input variables	Variable Meaning
NOT X1	No pains in eye	NOT X18	No swelling of eye
X1 and X2	Pains and redness of eye	X1 and X18	Pains and swelling of eye

Table 3. Some Additional variables for the Decision Tree.

The following rule sample sets are then obtained from the decision tree of Figure 13:

1. *IF* (X1 AND X2) and X18 and X20 and X8 and X11 and X17 *THEN* Cornel Ulcer
2. *IF* (X1 AND X2) and NOTX18 and X16 and X20 and X8 and X9 and X14 *THEN* Uveitis
3. *IF* (X1 AND X2) and NOTX18 and X16 and X6 and X5 and X7 *THEN* Glaucoma
4. *IF* NOTX1 and X8 and X10 and X5 and X12 *THEN* Muscular Degeneration
5. *IF* NOTX1 and X8 and X2 and X13 and X17 *THEN* Pink Eye

http://dx.doi.org/10.5772/51380

These rule sets are easily explain to means:

1. *IF* there is pains and redness of eye and swelling of eye and eye is sensitive to bright light and there is blurred vision and cloudy substance are formed in front of eye lens *THEN* the eye problem is *Cornel Ulcer*

2. *IF* there is pains and redness of eye and no swelling of eye and there are floater or flashes of light and eye is sensitive to bright light and there is blurred vision and the blurred vision improves with blinking of eye and there is discomfort after long concentrated use of eye *THEN* the eye problem is *Uveitis*

3. IF there is pains and redness of eye and no swelling of eye and there are floater or flashes of light and there are decrease in peripheral field of view and there is recorded family history of the eye problem and patient age is greater than 45 years *THEN* the eye problem is *Glaucoma*

4. *IF* there is no pains or redness of the eye and there is blurred and distorted vision and there is recorded history of the family history of the eye problem and slow recovery of vision after exposure to bright light *THEN* the eye problem is *Muscular Degeneration*

5. *IF* there is no pains of the eye and the eye is red and there is blurred vision and the eye is itchy or scratchy and there is watering discharge of the eye *THEN* the eye problem is *Pink Eye*.

Figure 13 shows an illustration of the extracted decision tree for some of the eye diseases.

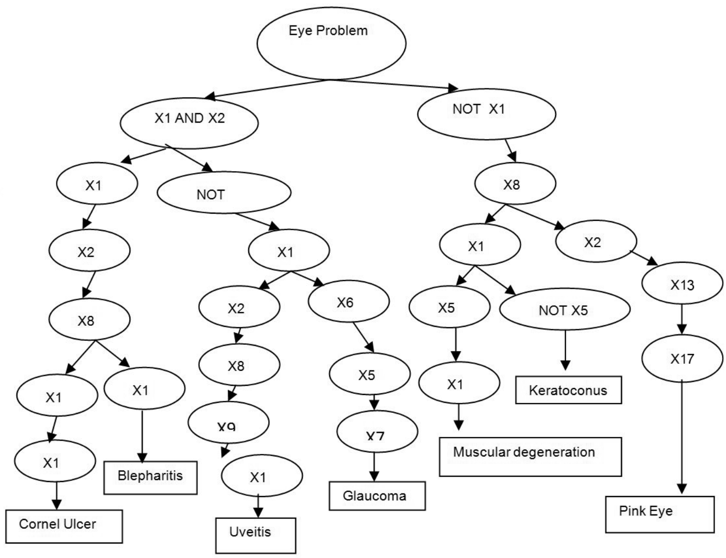

Figure 13. Decision tree for some the eye disease.

4.3. Performance Analysis of the system

To justify the performance of our diagnostic system, we conducted two analyses. The first is using a general performance scheme. Secondly, we carried out a number of tests at random using various physical eye examinations and patient's complain to see whether it agree with what it suppose to be.

4.3.1. Performance Benchmark

The proposed Neural Networks and Decision Tree for Eye Disease Diagnosis (NNDTEDD) architecture relies on a piece of software for easy eye disease diagnosis. The principles underlying diagnostic software are grounded in classical statistical decision theory. There are two sources that generate inputs to the diagnostic software: disease (H0) and no disease (H1). The goal of the diagnostic software is to classify each diagnostic as disease or no disease. Two types of errors can occur in this classification:

i. Classification of disease as normal (false negative); and

ii. Classification of a normal as disease (false positive).

We define:

Probability of detection $P_D = P_r$ (classify into H1 | H1 is true), or

Probability of false negative =1 - P_D.

Probability of false positive$P_F = P_r$ (classify into H1 | H0 is true).

Let the numerical values for the no disease (N) and disease (C) follow exponential distributions with parameters λ_N andλ_C, $\lambda_N > \lambda_C$, respectively. Then we can write the probability of detection P_D and probability of false positive P_F as

$$P_D = \int_t^\infty \lambda_C e^{-(\lambda_C x)} dx = e^{-\lambda_C t} \tag{7}$$

$$P_D = \int_t^\infty \lambda_N e^{-(\lambda_N x)} dx = e^{-\lambda_N t} \tag{8}$$

Thus P_D can be expressed as a function of P_F as

$$P_D = P_F^r \tag{9}$$

Where r = λ_F/λ_N is between 0 and 1.

Consequently, the quality profile of most detection software is characterized by a curve that relates its PD and PF, known as the receiver operating curve (ROC) [35]. ROC curve is a function that summarizes the possible performances of a detector. It visualizes the trade - off between false alarm rates and detection rates, thus facilitating the choice of a decision functions. Following the work done in [36], Figure 14 shows sample ROC curves for various values of r.

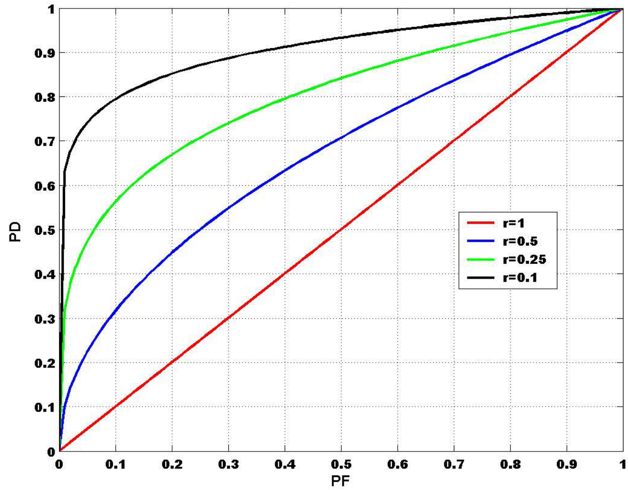

Figure 14. ROC curves for different r.

The smaller the value of r the steeper the ROC curve and hence the better the performance. The performance analysis of the NNDTEDD algorithms was carried out using MATLAB software package (MATLABR, 2009R) and the results compared with the collected data for cornel ulcer, uveitis, and glaucoma Figure 15, Figure 16 and Figure 17, respectively.

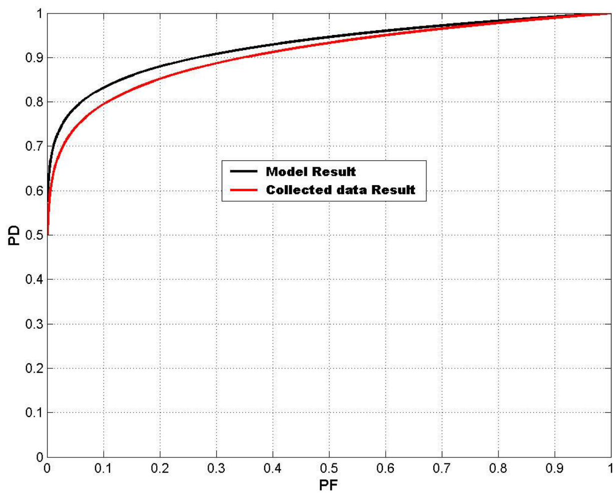

Figure 15. ROC curves for Cornel Ulcer diagnosis.

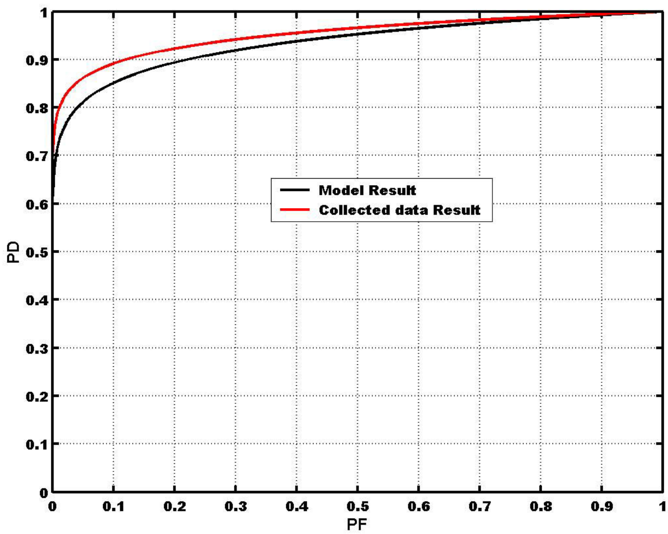

Figure 16. ROC curves for Uveities diagnosis.

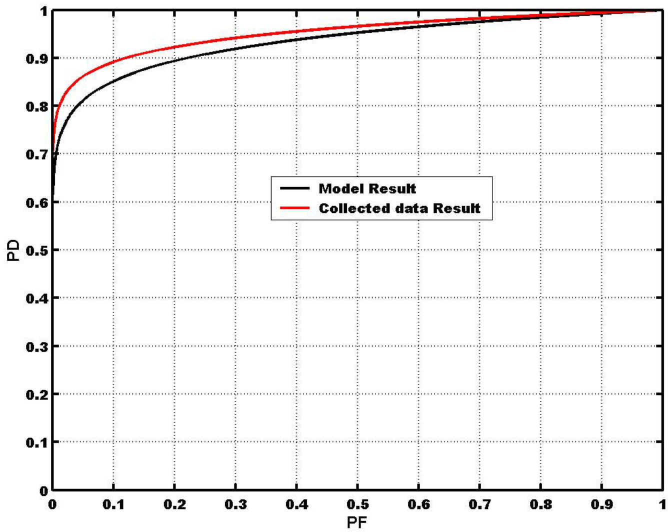

Figure 17. ROC curves for Glaucoma diagnosis.

http://dx.doi.org/10.5772/51380

4.3.2. Performance of the system using Random Tests

In testing the NNDTEDD, fifty different tests from the data sets for testing the system were carried out at random using various eye conditions and physical eye examinations combinations and the results compared with the expected result of the NNDTEDD. Where there was a match, success was recorded. In situations where there was no match failure were recorded. The total number of success = 46. Total number of failure = 4. Total number of test was 50.

$$Percentage\ succcess = \frac{46}{50} \times 100 = 92 \tag{10}$$

$$percentage\ failure = \frac{4}{50} \times 100 = 8\% \tag{11}$$

5. Conclusion

The research presented a framework for diagnosing eye diseases using Neural Networks and Decision Trees. This research extended common approaches of using a neural network or a decision tree alone in diagnosing eye diseases. We developed a hybrid model called Neural Networks Decision Trees Eye Disease Diagnosing System (NNDTEDDS). Neural networks have been successful in the diagnosis of eye diseases according to various symptoms and physical eye conditions. Decision trees have been useful in knowledge extraction from trained neural networks. They have been a means for knowledge discovery. We have obtained rules which explain the diagnosis of eye diseases according to various symptoms and physical eye conditions; these rules explain the knowledge acquired in neural networks by learning from previous samples of symptoms and physical eye conditions. The extracted rules can be used to explained how an eye disease is diagnosed hence removing the opacity in neural network alone. The extracted rules can also be used to train younger ophthalmologists. The proposed system was able to achieve a high level of success using the hybrid model of neural networks and decision tree technique. A success rate of 92% was achieved. This infers that combination of neural networks and decision tree technique is an effective and efficient method for implementing diagnostic problems.

6. Recommendations

This work is recommended to medical experts (ophthalmologists) as an aid in the decision making process and confirmation of suspected cases. Also, a non expert will still find the work useful in areas where prompt and swift actions are required for the diagnosis of a given eye disease covered in the system. Medical practitioners who operate in areas where there are no specialist (ophthalmologist) can also rely on the system for assistance.

Acknowledgements

We thank Dr. Monday Nkadam of the University of Port Harcourt Teaching Hospital for checking and correcting some of the symptoms that were used in the system. The authors also thank the staff of Linsolar and Odadiki eye clinic, Port Harcourt for using their data in training and testing the system.

Author details

L. G. Kabari[1*] and E. O. Nwachukwu[2]

*Address all correspondence to: ledisigiokkabari@yahoo.com

1 Department of Computer Science, Rivers State Polytechnic, Bori, Nigeria

2 Department of Computer Science, University of Port Harcourt, Port Harcourt, Nigeria

References

[1] Osheroff, J. A., Teich, J. M., & Middleton, B. F. (2006). A Roadmap for National Action on Clinical Decision Support. American Medical Informatics Association; 2006 June 13. Available at: http://www.amia.org/inside/initiatives/cds/. Accessed March 20, 2009.

[2] Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (2000). To err is human: building a safer health system. Washington, D.C.: National Academy Press.

[3] Miller, M., & Kearney, N. (2004). Guidelines For Clinical Practice: Development, Dissemination and Implementation. *International Journal of Nursing Studies*, 41(7), 813.

[4] Field, M. J., & Lohr, K. N. (2005). Clinical Practice Guidelines: Direction for a New Program. Institute of Medicine, Committee on Clinical Practice Guidelines. Washington, DC. National Academy Press.

[5] Musen, M. A. (1997). Modelling of Decision Support. *Handbook of medical informatics*, J. H. V. Bemmel and M. A. Musen, Eds. Houten: Bohn Stafleu Van Loghum.

[6] Bakpo, F. S., & Kabari, L. G. (2011). Diagnosing Skin Diseases using an Artificial Neural Network. *Artificial Neural Networks- Methodological Advances and Biomedical Applications, kenji suzuki (ed.)*, 978-9-53307-243-2, intech, Available from: http://www.intechopen.com/articles/show/title/diagnosing-skin-diseases-using-an-artificial-neural-network.

[7] Janghel, R. R., Shukla, A., Tiwari, R., & Tiwari, P. (2009). Clinical Decision support system for fetal Delivery using Ar tificial Neural Network. *2009 International Conference on New Trends in Information and Service Science*.

[8] Zhou, Q. (2009). A Clinical Decision Support System for Metabolism Synthesis. *2009 International Conference on Computational Intelligence and Natural Computing*.

[9] Robinson, A. J. (1994). An Application of Recurrent Nets to Phone Probability Estimation. *IEEE Transactions on Neural Networks*, 5(2), 298-305.

[10] Giles, C. L., Lawrence, S., & Tsoi, A. C. (1997). Rule Iinference for Financial Prediction using Recurrent Neural Networks. *Proceedings of the IEEE/IAFE Computational Intelligence for Financial Engineering*, New York City, USA, 253-259.

[11] Marakami, K., & Taguchi, H. (1991). Gesture Recognition Using Recurrent Neural Networks. *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through Technology, Louisiana, USA*, 237-242.

[12] Omlin, C. W., & Snyders, S. (2003). Inductive Bias Strength In Knowledge-Based Neural Networks: Application to Magnetic Resonance Spectroscopy of Breast Tissues. *Artificial Intelligence in Medicine*, 28(2).

[13] Chandra, R., & Omlin, C. W. (2007). Knowledge Discovery Using Artificial Neural Networks For A Conservation Biology Domain. *Proceedings of the 2007 International Conference on Data Mining, Las Vegas, USA*, In Press.

[14] Fu, L. (1994). Rule Generation from Neural Networks, IEEE Transactions on Systems. *Man and Cybernetics*, 24(8), 1114-1124.

[15] Niti, G., Anil, D., & Navin, R. (2007). Decision Support System For Heart Disease Diagnosis Using Neural Networks. *Delhi Business Review*, 8(1), January- June 2007.

[16] Lisboa, P. J. G., Ifeachor, E. C., & Szczepaniak, P. S. (2000). Artificial Neural Networks In Biomedicine. London: Springer-Verlag.

[17] Ledley, R. S., & Lusted, L. B. (1959). Reasoning Foundations of Medical Diagnosis. *Science*, 130, 9-21.

[18] Abbass, H. A. (2002). An Evolutionary Artificial Neural Networks Approach for Breast Cancer Diagnosis. *Artificial Intelligence in Medicine*, 25, 265-281.

[19] Zhou, Z. H., Jiang, Y., Yang, Y. B., & Chen, S. F. (2002). Lung Cancer Cell Identification Based On Artificial Neural Network Ensembles. *Artificial Intelligence in Medicine*, 24, 25-3.

[20] Vuckovic, A., Radivojevic, V., Chen, A. C. N., & Popovic, D. (2002). Automatic Recognition of Alertness and Drowsiness from EEG by an Artificial Neural Network. *Medical Engineering & Physics*, 24, 349-360.

[21] Mobley, B. A., Schechter, E., Moore, W. E., McKee, P. A., & Eichner, J. E. (2000). Predictions of Coronary Artery Stenosis by ANN. *Artificial Intelligence in Medicine*, 18, 187-203.

[22] Kara, S., Gu°ven, A., Okandan, M., & Dirgenali, F. Utilization of Artificial Neural Networks and Autoregressive Modeling in Diagnosing Mitral Valve Stenosis. (in press), *Computers in Biology and Medicine*.

[23] Wright, I. A., Gough, N. A. J., Rakebrandt, F., Wahab, M., & Woodcock, J. P. (1997). Neural Network Analysis of Doppler Ultrasound Blood Flow Signals: A pilot study. *Ultrasound in Medicine and Biology*, 23(5), 683-690.

[24] Salvi, M., Dazzi, D., Pelistri, I., Neri, F., & Wall, J. R. (2002). Classification and Prediction of the Progression of Thyroid-Associated Ophthalmopathy By An Artificial Neural Network. *Ophthalmology*, 109(9), 1703-8.

[25] Nguyen, H. T., Butler, M., Roychoudhry, A., Shannon, A. G., Flack, J., & Mitchell, P. (1996). Classification of Diabetic Retinopathy Using Neural Networks. *18th Annual International Conference of the IEEE Engineering In Medicine And Biology Society, Amsterdam*, 1548-1549.

[26] Tigges, P., Kathmann, N., & Engel, R. R. (1997). Identification of Input Variables for Feature Based Artificial Neural Networks-Saccade Detection in EOG Recordings. *International Journal of Medical Informatics*, 45, 175-184.

[27] Chan, B. C. B., Chan, F. H. Y., Lam, F. K., Lui, P. W., & Poon, P. W. F. (1997). Fast Detection of Venous Air Embolism in Doppler Heart Sound Using the Wavelet Transform. *IEEE Transactions on Biomedical Engineering*, 44(4), 237-245.

[28] Tu°rkog˘lu, I., Arslan, A., & I˙lkay, E. (2002). An Expert System for Diagnosis of the Heart Valve Diseases. *Expert Systems with Applications*, 23, 229-236.

[29] Schmitz, G. P. J., Aldrich, C., & Gouws, F. S. (1999). ANN-DT An Algorithm for Extraction of Decision Trees from artificial Neural Networks. *IEEE Transactions on Neural Networks*.

[30] Sethi, I. K. (1992). Layered Neural Net Design Through Decision Trees. *Proceedings of the IEEE*.

[31] Quilan, J. (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann, San Mateo, CA.

[32] Fabian, H. P., Chan, K. S., Ho, K. Y., & Leong, S. K. (2004). A Study on Decision Tree. *2nd Engineering & Technology Student's Congress. Kota Kinabalu.: SKTM.*

[33] Haykin, S. (1999). Neural Networks: a Comprehensive Foundation. Second Edition. Prentice Hall., 842.

[34] Yedjour, D., Yedjour, H., & Benyettou, A. (2011). Explaining Results Of Artificial Neural Networks. *Journel Of Applied Scinces*, 2(3).

[35] Trees, H. V. (2001). Detection, Estimation and Modulation Theory- Part I. John Wiley, New York.

[36] Huseyin, C., & Srinivasan, R. (2004). Configuration of Detection Software: a Comparison of Decision and Game Theory Approaches. *Decision Analysis*, 1(3), 131-148.