Neural Approach Applying for Soil Classification System

Khin Mo Sue, Tin Mar Kyi
Computer University, Loikaw
khinmosue.ktms@gmail.com, tinmarkyi@gmail.com

ABSTRACT

*The main aim of this paper is to develop a system for classifying the soil class of Myanmar by the help of Error Back-propagation algorithm (EBP) which is the most widely used algorithm among Artificial Neural Network (ANN) technique. This system, "**Soil Classifier**" includes two parts in general, training and testing. During the training phase, the Soil Classifier accepts nine inputs. These inputs are soil type, land use type, land form, soil depth, soil texture, soil PH and the percentages of each of three types of plant nutrients. The plant nutrients are Nitrogen (N), Phosphorus (P) and Potassium (K). After accepting the nine inputs, the Soil Classifier will generate one of three types of outputs whether the soil is good class, fair class or poor class. The Soil Classifier uses Multi Layer Feed-forward Neural Network.*

Keywords: Error Back-propagation algorithm, Artificial Neural Network (ANN), Multilayer Feed-forward Neural Network.

1. Introduction

Soil is the most important medium for plant growth. Understanding the characteristic of soil, in particular, knowing how to maintain good quality of soil is important for plantation. To know how to improve difficult soil is importance for successful agriculture. The classification of soil class is such a time consuming work and needs many physical tests and it is difficult to determine. Because there is no certain fact to point out the soil class whether it is good, fair or poor when calculate. The nature of soil data is in the form of non-linear.

In this proposed system, the classification of soil class is employed by the help of Artificial Neural Network. Supervised learning Back-propagation algorithm is used to solve this problem.

The main aim if this paper is to apply the fundamental concepts of Artificial Neural Networks effectively and to know about the nature of neural algorithm and theory. The aim of employing the flexibilities of computer technology and developing the classification of soil types by using error back propagation technique are also included. Another aim is to provide the agriculture by knowing the soil type instead of performing in such a time consuming work by using this proposed system.

The proposed system can classify soil class of all location of Myanmar. The soil classifier trained soil data that can be found in Myanmar.

2. Related Works

According to the modern classification, there are three main types of soil classes being recognized in the Union of Myanmar. The characteristics of these soils are determined mainly upon four facts, first, the physical and mineral composition of the parent material, second, the relief (physical features), third the climate under which the soil material has been developed and the last is the vegetation. Soil classification has generally been made on the basis of the distribution of the important land resources for agriculture [2]. J.M. Zurada is presented for multilayer feed forward neural network [1]. R.Collan is presented for Back-propagation algorithm. Twe Ta Oo [5] is presented or classification of soil class and its paper is implemented for sentence composition. R.Collan [3] is presented for the learning rate and error rate of the proposed system. Symon Hykin [4] is presented for activation functions used in this proposed system.

3. Multilayer Feedforward Neural Network

A multilayer feed-forward neural network consists of an input layer, one or more hidden layers, and an output layer. Each layer is made up of units or nodes. The inputs to the network correspond to the attributes measured for each training tuple.

The inputs are fed simultaneously into the units making up the input layer. These inputs pass through the input layer and are then weighted and fed simultaneously to the second layer of "neuron like" units, known as a hidden layer. The outputs of the hidden layer units can be input to another hidden layer, and so on. The number of **hidden** layers is arbitrary, although in practice, usually only one is used. The weighted outputs of the last hidden layer are input to units making up the output layer, which emits the network's prediction for given tuples.

The units in the input layer are called input units. The units in the hidden layers and output layer are sometimes referred to as neurodes or output units.

3.1 Procedure of Multilayer Feedforward Neural Network

The general procedures of multilayer feed-forward neural network are as follow-

1. First apply the inputs to the network and work out the output. The initial output could be anything, as the weights were small random numbers.
2. Next, work out the error for hidden neuron. The error signal is the difference between desire output and actual output of the system. In other words:

   Error = Output (1 - Output) (Target - Output)
3. Change the weight.
4. Calculate the Errors for the hidden layer neurons. Unlike the output layer, the error of hidden neuron can not be calculated directly as output layer neuron because it does not have target output. This is done by taking the Errors from the output neurons and running them back through the weights to get the hidden layer errors.

   $Error_H = Output_H (1 – Output_H) \Sigma Error_O W_{OH}$
5. Having obtained the error for the hidden layer neurons now process as in stage 3 to change the hidden layer weights. By repeating this method a network of any number of layers can be trained.

3.2 The Back-propagation Algorithm

The most well-known algorithm in Artificial Neural Network field is Error Back-propagation algorithm (EBP). The Error Back-propagation algorithm consists of two sweeps of the network, first a forward sweep from the input layer to the output layer. Second, the backward sweeps from the output layer to the input layer. A unit in the hidden layer will send activation to every unit in the output layer during the forward sweep and so during the backward sweep a unit in the hidden layer will receive error signals from every unit in the output layer.

4. The Procedures of Soil Classifier

Step1: Accept soil data as input of the system.
Step2: Associate pattern (a pair of input pattern with desire output pattern).
Step3: Initialize the weight value by taking small random value.
Step4: Calculate the total net input (**net$_j$**) for each neurons of second layer or hidden layer.

$$\text{netj} = w0 + \sum_{i=1}^{n} xiwij$$

Where **x_i** is input value and **w_{ij}** weight value.
Step5: Calculate the output oj for each neuron of hidden layer. The activation function of output is sigmoid function.

$$oj = \frac{1}{(1+\exp(-netj))}$$

Step6: Calculate the total net input for each neuron of third layer or output layer.

$$\text{netj} = w0 + \sum_{i=1}^{n} xiwij$$

Step7: Calculate the output for each neuron of output layer.

$$oj = \frac{1}{(1+\exp(-netj))}$$

Step8: Compute error δj for each output neuron.

$$\delta j = oj(1-oj)(tj-oj)$$

Where **o_j** is actual output of the network and **t_j** is the target output of the network.
Step9: Compare the actual error with pre-determine error whether the actual error is within the tolerance rate. If the actual error is not within tolerance, then this error signal is back propagate to adapt the weight values of the network.
Step10: The new weight, **W^+_{new}** value is calculated

$$\mathbf{W^+_{new}=W_{old}+\eta w_{ij}}$$

where **W_{old}** is the old weight value of the network.
Step11: After adapting the weight values that connect between output and hidden layer, the error for hidden layer's neurons is calculated. Hidden layer has no target value so the error of output neuron is used to back propagate.

$$\delta j = (1-oj)\sum_{k} \delta kwkj$$

This step takes repeat action until the error of last neuron of hidden layer is calculated.

Step12: The weights of all layers are adapted.

$$\Delta wij(n+1) = \eta(\delta joi) + \Delta wij(n)$$

5. The Network of Soil Classifier

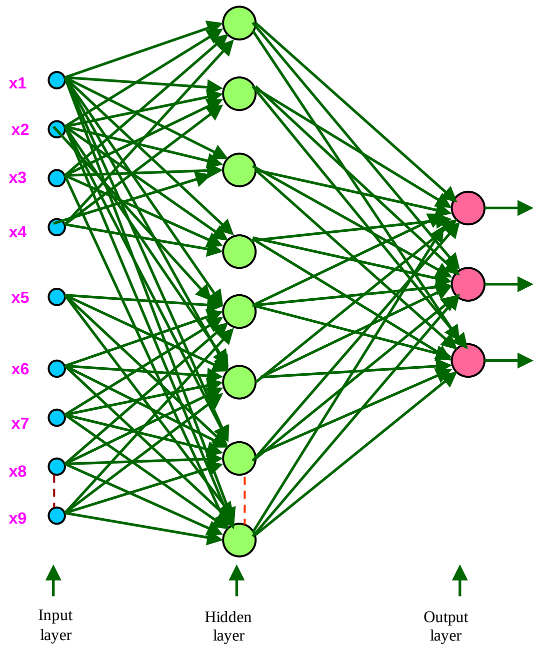

Figure1. The multilayer feed-forward neural network of soil classification system (9-20-3 architecture)

5.1 The Architecture of the Soil Classifier

The architecture of this proposed system includes three layers. The first layer of system is input layer, hidden layer for second layer and the last one is output layer. The input layer consists of nine nodes to accept nine kinds of soil data. The first node of this layer accepts the soil type. This soil types vary in nineteen kinds.

The second node accepts one of twelve kinds of land use type. The third node accepts one of twenty five kinds of Land form. The fourth node accepts soil depth. The fifth node accepts amount of nitrogen. The sixth node accepts phosphorus. The seventh node accepts amount of potassium. The eighth node accepts soil ph and the ninth node accepts one of eleven kinds of soil texture.

After accepting these inputs the system will generate point values for each input to propagate to the network. Three types of activation functions are used to calculate output for each layer.

For the input layer, every node use identity activation function to calculate net output. Because the input layer has only the task that is to propagate input value as the same value as it accepted. Then these input values are multiplied with small random weight values and summed as total net input of hidden layer neurons. This step use binary threshold function.

For the output of each neuron of both hidden and output layer neurons use sigmoid function. The sigmoid function usually produces the values between 0 and 1. There are two types of sigmoid functions, bipolar sigmoid function and hyperbolic tangent function.

In this proposed system the bipolar sigmoid function is used. After calculating the output values of neurons that lie on the final output layer, these output values is compared with the desire output or target output whether the error is within tolerance value.

Error rate of this system is defined in the range between 0.0 and 0.3. The error rate of the system is inversely proportional to the training count. The smaller the error rate value the longer training time will be. The optimum value of error rate value for this system is 0.002.

There are three types of target pattern to represent good soil class, fair soil class and poor soil class. The standard patterns are 100 for good, 010 for fair class and 001 for poor class. If not then the error signal will back propagated and the new weights are computed.

The soil classifier used learning rate parameter to calculate weight change between different layer units. Another purpose of using learning rate parameter is to stabilize the network during training phase. There is no default value for learning rate parameter because this value may vary, depends on the system that is used. In this proposed system the optimum value that helps to stabilize the network is 0.25. This step will iterate until the error is minimized

One of the importance facts of determining the architecture of a Neural Network is the number of hidden layer and hidden neurons. Because too many numbers of hidden neurons make the network be error prone but too lass of these make the network make the network be difficult to converge the target result. The number of hidden neurons used in this proposed system is 20 neurons. Only one layer of hidden layer is sufficient for most system. This proposed system used only one hidden layer.

The sample data set of the soil classifier is as shown in the following table.

Table1. The sample data set of soil classifier

ST	LUT	LF	SD	STX	SPH	N	P	K	C
0.1	0.1	0.1	0.12	0.1	0.1	0.1	0.1	0.1	100
0.1	0.1	0.1	0.1	0.2	0.3	0.2	0.3	0.1	100
0.2	0.2	0.1	0.1	0.2	0.3	0.2	0.3	0.1	010
0.2	0.3	0.3	0.1	0.5	0.3	0.2	0.2	0.2	010
0.2	0.8	0.5	0.1	0.7	0.2	0.2	0.3	0.2	100
0.2	0.8	0.5	0.1	0.7	0.3	0.2	0.3	0.2	100
0.2	0.2	0.1	0.1	0.2	0.3	0.3	0.2	0.1	100
0.5	0.4	0.2	0.2	0.3	0.3	0.2	0.3	0.2	010
0.8	0.8	0.1	0.2	0.6	0.2	0.3	0.3	0.1	001
0.8	0.3	0.1	0.2	0.2	0.2	0.3	0.3	0.2	001
0.8	0.3	0.1	0.1	0.2	0.2	0.1	0.3	0.1	001
0.7	0.3	0.2	0.1	0.2	0.1	0.2	0.2	0.1	001

ST	Soil Type	SPH	Soil PH
LUT	Land Use Type	N	Nitrogen
LF	Land Form	P	Phosphorus
SD	Soil Depth	K	Potassium
STX	Soil Texture	C	Class

The following table shows three types of output class of the proposed system.

Table2. Output of Soil Classifier

No	Class Name	Output Pattern
1	Good	100
2	Fair	010
3	Poor	001

The following table shows the nine kinds of soil attributes and their range that used in this proposed system.

Table3. The range of soil data

No	Attributes Name	Range
1	Soil type	0.1 to 0.21
2	Land use type	0.1 to 0.13
3	Land form	0.1 to 0.28
4	Soil depth	0.1 to 0.3
5	Soil texture	0.1 to 0.12
6	Soil ph	0.1 to 0.3
7	Nitrogen	0.1 to 0.3
8	Phosphorus	0.1 to 0.3
9	Potassium	0.1 to 0.3

5.2 System Flow Diagram

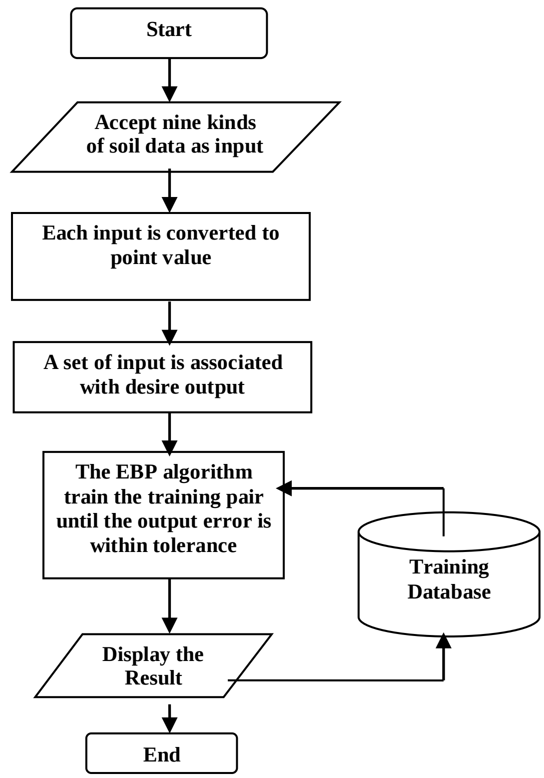

Figure2. System flow diagram of soil classifier

5.3 The System Implementation

The following figures represent the training pattern that is identified by user. The upper portion of the figure3 shows the input pattern and the lower portion shows the output pattern that defined by user.

Figure3. Data inputs for training.

The following figures show the converge condition of the training data. Three nodes of output layer represent the desire output that has been identified by user.

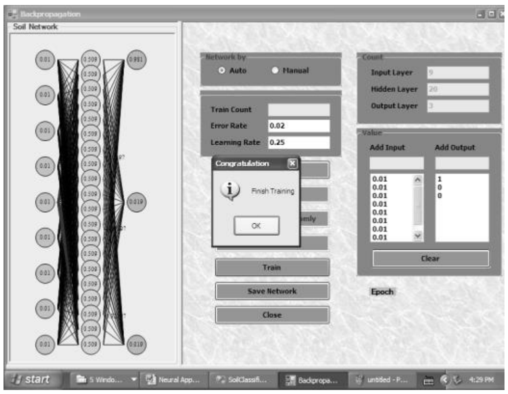

Figure4. The result of training pattern.

6. Conclusion

Myanmar is an agriculture country. Agriculture is very essential and fundamental for both economy and everyday life of Myanmar. To be a successful agriculture it is very important to determine the soil class because soil is the most important thing for plant growth. It takes many years to become an inch of plantable soil layer. So to maintain the quality of soil is also very importance. It is also importance to determine which kind of plant should be planted. By selecting correctly the suitable kind of crop the farmer can get good profit. Soil Classifier is implemented to fulfill these necessary by using Artificial Neural Network concepts. The most well known algorithm of ANN, Error Back-propagation Algorithm is used to develop this system. It helps the soil classifier to learn many examples based on its nature of supervised learning. This system only accepts the soil type and characteristic of Myanmar.

References

[1] J.M. Zurada, "Introduction to Artificial Neural Systems", Info Access Distribution Pte Ltd, 1992.

[2] Ministry of Agriculture and Irrigation, "Soil Type and Characteristic of Myanmar" , press by Ministry of Agriculture and Irrigation, 2009.

[3] R.Collan, "The Essence of Neural Networks", Prentice Hall, 1990.

[4] Symon Hykin, "Neural Network: A Comprehensive Foundation", Second Edition, Prentice Hall, 1999.

[5] Twe Ta Oo, "Classifying Sentence Composition Using Artificial Neural Networks", Pro of the conference on Applied Information and Communication Technology, September.2009.