International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 5 Issue: 1
ISSN: 2454-4248
57 – 60

Artificial Neural Networks for Soil Quality and Crop Yield Prediction using Machine Learning

Dr. T. Venkat Narayana Rao
Professor, CSE, SNIST
Sreenidhi Institute of Science and Technology
Hyderabad, India

S. Manasa
Student , CSE, SNIST
Sreenidhi Institute of Science and Technology
Hyderabad, India

Abstract: Agriculture is the main stream on which farmers depend. Many surveys have proved that suicide rate of farmers is proliferate over years. The main reasons for the increase in suicide rate are weather conditions, debts, lack of details about the soil. In some remote areas farmers lack information about soil quality, soil nutrients, soil composition and may choose wrong crop to sow which results in less yield. So as to overcome the issues faced by farmers we are trying to implement a model using Artificial Neural Networks (ANN) which predicts the soil quality taking input as several important parameters related to soil. This paper mainly focuses on predicting the crop yield using the ANN which is completely a software solution and also recommends suitable fertilizers to gain high yield of crops.

Keywords: *Artificial Neural Networks (ANN), Backpropagation Network, soil quality, crops, pH values, fertilizers*

*****

I. Introduction

Agriculture sector plays a major role in Indian economy, as 70 percent households in India depends purely on this field [1]. In 2014, the National Crime Records Bureau of India reported 5,650 farmer suicides [2]. We have several hardware devices like pH meter to test the pH value of soil but it is a time consuming process, expensive and also cannot assure accurate results. Also not all farmers are literates and may not know the amount of fertilizers to be used for a fertile land. Analysis of soil condition is one of the primary tasks prior to sowing a crop. Although a farmer's experience from past results might help him in estimating the soil quality and crop yield to some extent but this type of estimation is not always accurate. The evolution of technology has brought many changes in agriculture. Now we are trying to provide a idea to improve the agricultural sector by using Machine Learning. Efficient techniques can be developed for solving complex soil data sets using machine learning to improve the effectiveness and accuracy of the Classification of large soil data sets [3].The idea of Machine Learning for crop yield prediction is completely a software solution and guarantees to give accurate results in fraction of seconds. In this paper we are focusing on predicting soil quality and crop yield using neural networks.

Artificial Neural Networks (ANN)

An Artificial Neural Network (ANN) is a computational model which resembles the structure and functions of biological neural networks. The flow of information through the network will affect the ANN structure. This is because a network first learns and then changes depending upon the input and output i.e. depends on the data given in the dataset. In simple terms we call it as neural network. In biological terms an ANN functions similar to brain. In the brain, an electrical signal is sent by a neuron through axon which splits into many branches. There is an area called a synapse at the end of each branch. An ANN reminds us of the functioning of the brain because there are weighted connections which are called as synapses between simulated neurons where signals it receives (numbers) are summed and then (with most neuron models) a signal is sent if a certain threshold is reached [4].

Back propagation

The study says that data is fed to a Back Propagation Network (BPN) to evaluate the test data set [5]. The basic usage of back propagation is to improve the accuracy of predictions in machine learning. Back propagation is a common method of training artificial neural networks so as to make the model be prepared to compute any kind of task. Back propagation algorithm is a type of supervised machine learning algorithm which requires input and output to be known. We build a neural network model generally using back propagation algorithm. The procedure is desired outputs are compared to achieve system output and then the systems are tuned by adjusting the weights so that the accuracy is high. BPN is able to enhance the ability to predict and competition between different species under changed environmental conditions and their potential applicability in other crops to grow as well [6].

At the start of the training process weights are assigned randomly. So as to acquire high accuracy the process of training the data must be iterative. The part of data chosen

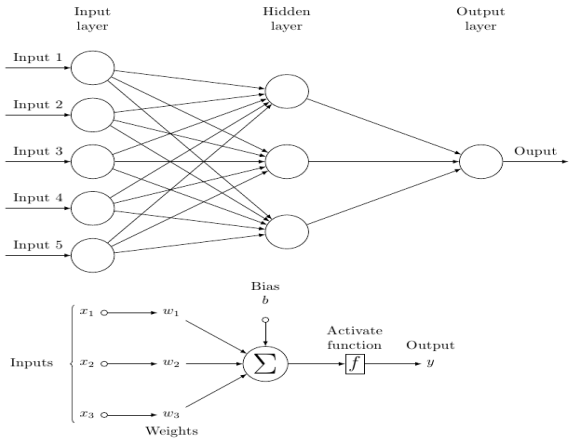
 ISSN: 2454-4248
57 – 60

for training is given to the neural network model iteratively. The desired output is compared with the system's obtained output. The difference is nothing but the error. This error is back propagated to the network model and weights are adjusted so as to reduce errors. As this process is an iterative process the adjustment of weight continues until the error is acceptable. In terms of accuracy this process continues until we achieve desirable accuracy. The system proposed tries to overcome these drawbacks and predicts crops by analyzing structured data [7].

Most of the existing systems are hardware based and must be maintained in a proper manner. pH meter is a hardware device which is used to test the soil and determine the alkalinity. Sensors which are expensive are used to determine various soil properties. Taking all of these into consideration we felt there is a need for software to determine soil properties like alkalinity, pH values without tedious work. Before we determine crop yield we must check the soil quality, determine nutrients in the soil, predict which crop is to be sown, recommend suitable fertilizers. All these tasks can be done by using Machine Learning (ML). The only requirement for this software is the datasets which contain lots of statistical structured data related to our problem. Being a complete software solution we can assure accurate results without time lag and maintenance is not required.

II. Architecture

Neural networks are represented in layers form in which each layer consists of a number of nodes depending on the number of input parameters. Hidden layer can contain any number of nodes as per our requirement. Output layer also contains the nodes depending on the number of output parameters. The input layer takes the information or data from outside world. Input layers do not perform any processing or computations. They just pass data to the hidden layers. Hidden layers are like abstraction layers which perform computations on the data received from input layer and transfer the result to output layer. The output layers just give output to the external world. The hidden layers are used to process the input data using activation function and produce the output. Activation function decides whether a neuron should be activated or not by calculating the weighted sum and further adding bias with it [8]. If the sum of weighted sum and bias is greater than a threshold value then the neuron is activated or fired. Weighted sum is the summation of product of weights and inputs.



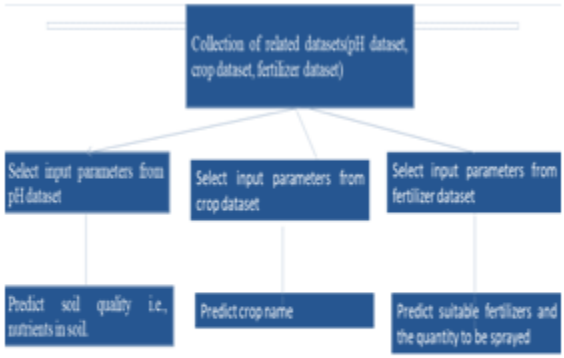
2.1 Representation of Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) consists of three layers namely input layer, hidden layer and output layer. A neural network model can have more than one hidden layer. The accuracy would be high if the number of hidden layers is high. The basic representation of neural networks is shown in the fig. 2.1. We are using backpropagation method to build and train a neural network model. The training of the Backpropagation network (BPN) is done in three stages- the feed forward of the input training pattern, the calculation and backpropagation of the error and updation of weights [9]. The main aim of the neural network model is to train the model to achieve a balance between the ability of a model to respond and its ability to give reasonable responses to the input that is similar but not identical to the one that is used in training. In simple terms training a model by giving a number of patterns and then test the model using a new data which is similar to that present in the dataset but not identical.

We will add hidden layers one by one using dense function.

- output_dim: It is simply the number of nodes you want to add to the output layer.
- In Neural Network we need to assign weights to each mode which is nothing but importance of that node.
- At the time of initialization, weights should be close to 0 and we will randomly initialize weights using uniform function.
- input_dim: This parameter is needed only for first layer as model doesn't know the number of our input variables.

 ISSN: 2454-4248
57 – 60



2.2 Architecture of proposed system

The architecture shown in the fig.2.2 gives a brief of what is to be predicted using which dataset. We need three datasets – pH dataset, crop dataset, fertilizer dataset. pH dataset contains pH values of soil, pH rate, type of soil, Nitrogen (N), Phosphorous (P), and Potassium (K) content in the soil, Zinc (Zn), Iron (Fe), Manganese (Mn), Boron (B) content in soil, maximum temperature and minimum temperature. By using pH dataset we predict the soil quality by which we mean nutrient content in the soil is predicted given pH of the soil, type of soil, nutrient content, pH rate, maximum temperature and minimum temperature as input. Crop dataset contains crop names and nutrients required for a crop to give high yield. By using crop dataset we predict the crop names given nutrient content as input. Fertilizer dataset contains crop names, fertilizers required to have high crop yield and quantity of fertilizers to be sprayed. By using fertilizer dataset we predict the fertilizers and their quantities required to have a high crop yield given crop name as input.

III. Machine learning for prediction of crop yield

We have numerous algorithms in Machine learning which can be used to predict the soil quality and crop yield. But we have chosen Backpropagation algorithm to predict the soil quality because of its high accuracy compared to other algorithms. We can take as many input parameters to get more accurate output. We consider taking pH values of soil, type of soil, nutrients content in the soil, pH rate, maximum temperature and minimum temperature as the input parameters. We mainly concentrate on taking Nitrogen (N), Phosphorous (P), and Potassium (K) as nutrients of soil. Zinc (Zn), Iron (Fe), Manganese (Mn), Boron (B) content in soil can also be taken as another input parameter to gain even more accuracy.

Backpropagation algorithm is a method used in artificial neural networks commonly to train deep neural networks. In the context of learning, backpropagation is commonly used by the gradient descent optimization algorithm to adjust the weight of neurons by calculating the gradient of the loss function [10]. The loss function gives us the difference between the desired output and system output. Based on this difference we can propagate back to reduce the loss function. The dataset must contain columns including the attributes pH values of soil, pH rate, type of soil, Nitrogen (N), Phosphorous (P), and Potassium (K) content in the soil, Zinc (Zn), Iron (Fe), Manganese (Mn), Boron (B) content in soil, maximum temperature and minimum temperature. After the dataset is learnt by the model, the model will be able to give us the output nutrient content (N, P, K) in the soil when given pH as input.

Taking pH values of soil, type of soil, pH rate, maximum temperature and minimum temperature as input parameters preprocess the dataset and prepare a structured dataset which machine can understand. The output to be determined is the nutrient content (N, P, K) in the soil. Spilt the dataset into two parts train data and test data. We consider about 70% of the data as train data and 30% of the data as test data. Build a neural network model using backpropagation method taking input layers depending upon the number of input parameters, number of hidden layers of our choice, and output layer. Any number of hidden layers can be taken. More the number of hidden layers greater is the accuracy.

There is another phase of the problem which includes predicting the crop to be sown in particular land or soil. For this take crop name, nutrients required for a crop to give high yield in another dataset. Take nutrients as input parameters and crop name as output parameter which is to be predicted. Preprocess this dataset and split the dataset into two parts train data and test data. Similar to the above model build another model and from this get the crop name as output.

Once the crop to sown is known then recommending the fertilizers is to be done. Take another dataset which has information about name of the crop, its respective fertilizers requirement, quantity of fertilizers to be sprayed. Consider input as crop name, build a neural network model and obtain the fertilizers required along with its quantity to be sprayed. In this manner a neural network model can serve the purpose of farmers by predicting the soil quality and crop yield. We can also consider taking soil moisture content, average rainfall as input parameters in predicting the soil quality. As we increase the number of intake of parameters the accuracy also increases and the prediction will be efficient.

IV. Applications

The perspective of this paper is to focus only on agriculture and farming. As we are advancing in technology day by day there is a need of using the technologies to get the problems faced by farmers done right. One such experiment is the idea presented above. The idea proposed mainly focuses on two things

 ISSN: 2454-4248
57 – 60

4.1 Predicting soil quality

Analyzing soil quality is the primary task in agriculture because it is the root for further steps we take. Considering the soil conditions and weather conditions as input parameters we are trying to predict the nutrient content (N, P, and K) in the soil. Next step is using the nutrient contents in the soil as input and predicting a suitable crop for the soil. With this idea farmers need not struggle to decide which crop to grow. Therefore we can save time, completely avoid the usage of hardware and need not consider maintenance factor because it is completely a software solution.

4.2 Recommending fertilizers

Once we know the crop to be sown in a particular land then further step is to improve the crop yield. This is done by using suitable fertilizers for the crop. We can use a neural network model to recommend the fertilizers. Usually farmers depend on people who sell fertilizers to recommend fertilizers suitable for their crops which cannot be trusted always. By using this idea we can recommend correct fertilizers and also their quantities accurately.

As this paper mainly focuses on overcoming the problems faced by farmers it is completely related to agriculture and the applications of this idea focus only on agriculture. We can also take more number of parameters than those mentioned in this paper. As the number of inputs increase the results we obtain would be very accurate. In this paper we proposed three neural network models which takes different parameters and give us accurate results.

V. Conclusion

In this paper we have proposed a system that uses Neural Networks to predict the soil quality and crop yield. The reason behind using neural networks is for more accurate results. This system can reduce the difficulties faced by farmers to some extent by helping them to choose the crop to be sown depending on the soil quality. The system can be further enhanced to add the functionality of detection of crop diseases by using image processing. For this we need to collect the images of diseased crops and train the machine to detect any kind of disease by using image processing. The idea presented in this paper can help increase the economy of agriculture sector and reduce the suicide rate of farmers.

References:

[1] Food and Agricultural Organization. (2006). The state of Agricultural Commodity Markets. 37-39

[2] National Crime Reports Bureau, ADSI Report Annual – 2014 Government of India, p. 242, table 2.11

[3] A. Kumar & N. Kannathasan, (2011), “A Survey on Data Mining and Pattern Recognition Techniques for Soil Data Mining “, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3.

[4] B.A. Smith et al, Artificial Neural Networks for Automated Year round Temperature Prediction. Computers and Electronics in Agriculture 68 (2009) 52–6

[5] Shivnath Ghosh,Santanu Koley, “Machine Learning for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks” IJRITCC, vol. 2, Issue 2,292-297,2014

[6] W. Y. Chen, J. Yang, Z. Y. Chen, F. Y. Luo, C. L. Wu and Z. S. Peng, “Predicting photosynthetic rate of sunflowers using back propagation neural network based on uniform design”, African Journal of Agricultural Research Vol. 6(27), pp. 5817-5821, 19 November, 2011.

[7] Echochem Online. (2009). Soil Health and Crop yields. Last modified January 28th 2009.Retrieved on March 4th 2009 from http://ecochem.com/healthy_soil.html

[8] http://www.geeksforgeeks.org/activation-functions-neural-netwoks/

[9] "What is Backpropagation?". deepai.org. Nielsen, Michael A. (2015). "Chapter 6". Neural Networks and Deep Learning.

[10] Saeid Iranmanesh, M. Amin Mahdavi, “A Differential Adaptive Learning Rate Method for Back-Propagation Neural Networks”, World Academy of Science, Engineering and Technology 26, 2009.