



SOIL QUALITY PREDICTION FOR DETERMINING SOIL FERTILITY IN BHIMTAL BLOCK OF UTTARAKHAND (INDIA) USING MACHINE LEARNING

JANMEJAY PANT^{1,*}, PUSHPA PANT², R. P. PANT¹, ASHUTOSH BHATT³, DURGESH PANT⁴, AMIT JUYAL⁵

¹*Graphic Era Hill University, Bhimtal, India*

²*M.B. Govt. P.G College, Haldwani, India*

³*Birla Institute of Applied Sciences, Bhimtal, India*

⁴*Uttarakhand Open University, Haldwani, India*

⁵*Graphic Era Deemed to be University, Dehradun, India*

**Corresponding author: geujay2010@gmail.com*

ABSTRACT. Agriculture plays a vital role in the Indian economy. The growth of agriculture sector is based on the type of gift we have got from the nature. It varies state to state, district to district, taluka to taluka, block to block and even village to village. This study is confined to Bhimtal block of Nainital district. The main purpose of agriculture is growing crops and raising livestock. In order to grow the crops several types of agri-inputs are required, among them fertile lands have the great significance in crop cultivation. As far as fertile land is concerned it solely depends on the quality of the soil in terms of producing the nutrients for the crops. The available nutrients present in soil can be evaluated and measured by soil testing tools. The appropriate quantity of soil nutrients supplied to the soil can also be determined by this tool. The quantity of supplied nutrients is based on soil fertility and crop needs. In this study we have classified different soil features such as OC (Organic Carbon), P (Phosphorus), K (Potassium), Mn (Magnesium) and B (Boron). In order to make meaningful inferences and estimates, machine learning techniques especially ANN network with two activation functions relu and tanh are used in this study. For categorizations and predictions we have used village wise soil test report values. This kind of practice will not only help stakeholders

Received September 11th, 2020; accepted October 6th, 2020; published December 11th, 2020.

2010 *Mathematics Subject Classification*. 47N10.

Key words and phrases. machine learning; soil fertility; classification and prediction; activation function; keras; tensorflow.

©2021 Authors retain the copyrights of their papers, and all open access articles are distributed under the terms of the Creative Commons Attribution License.

to mitigate the expenditure of continuously supplying fertilizers to soil but it would also be cost effective, less time consuming and more profitable for stakeholders. In this regard data was compiled, classified, tabulated, presented, analyzed and it can be seen that relu activation function has ensured higher accuracy over tanh activation function. It is expedient and necessary to mention here that out of the five classified soil nutrient parameters relu activation function has shown better performance in respect of four classified soil nutrient parameters while tanh gave better performance in only one classified soil nutrient parameter.

1. INTRODUCTION

Agriculture is one of the biggest sectors in India that affects the Indian economy but due to urbanization and industrialization the cultivable land is reducing and soil fertility is declining. Therefore it becomes a challenge to boost the agricultural production without harming the environment. This challenge can be accomplished by increasing the fertility of the soil through supplying the fundamental nutrients to the plant at the right time in an acceptable amount. Soil management is also necessary for increasing the crop production and this can be done by improving and maintaining the soil nutrients. Crop production can be enhanced by selecting a relevant soil management strategy. It becomes easy for the experts to make predictions and take decisions on the suitable soil resource management if soil problems and crop related issues get identified at the right time. In the current scenario, Machine learning techniques are capable to solve the classification and prediction problems in agriculture. The challenges faced by soil scientists and domain experts in the terms of crop and soil management may definitely minimized by ML models. Low quality soil always mitigates the crop health and production. Instability in soil has always been responsible and significant factor for the crop health, crop yield and increment of crop production [1]. It has been stated earlier that reducing the soil fertility level and improper use of soil nutrients may cause food crisis for the world [2]. Many studies have been carried out towards in agriculture for dealing with soil problems using machine learning but no comprehensive study has been conducted so far. Soil fertility can be defined as ability of soil to provide the required nutrients and water to the crop or plant for their growth. In this regard a review paper [3] was published to discuss the methods which can be used for fertility prediction. This paper [4] found that SVM gives better results over the neural networks to provide PTF (Pedotransfer Function). Machine learning in soil management and crop production is an emerging field of research in the recent time. A lot of researches are going on this direction. In earlier days, Support vector Machine is used to estimate soil properties values and soil type classification based on the physiochemical properties soil [5]. A comprehensive SVM (Support Vector Machine) based classification [6] was implemented for urban soil quality assessment. The main aim of the review paper [7] is to encapsulate all the soft computing techniques in the field of agriculture and biological engineering. Different machine learning techniques were applied to analyze the soil texture in southwest china in this paper [8]. For this purpose SVM and ANN are used and

finally these techniques were compared in terms of accuracy. The paper [9] proposes a series of models to evaluate the soil nutrients. In this paper Multiple Linear Regression (MLR), a statistical approach, is also used along with SVM and ANNs to train the model. The results demonstrate the prediction accuracy of each model and then compare the accuracy. Some research has also done to predict the contents of Organic matter and the pH of paddy fields [10]. Various models have been used to predict soil properties over the last few years. Soil phosphorus is predicted using the statistical techniques [11] in this paper they tried to develop best model to predict phosphorus (P) using the different statistical techniques which have intelligent methods and regression models [11]. This review paper [12] elaborates the applications of ML methods in agriculture. In the previous study [3] Soil fertility problems were identified and predicted by ML. This paper [13] explained the j48, K-nearest neighbors (KNN) to classify and predict the wheat yield. In this paper [14] the soil parameters are classified by different classifiers of various families such as decision tree, support vector machine, random forest in India. Regression methods were applied in this paper [15] to predict soil fertility for various available nutrients in Maharashtra. In this paper [16] soil nutrients and pH classification and Soil features predictions was described by Extreme Learning Machine. This work was done in the state of Kerala. Machine learning techniques are also used for quantified the features as very low, low, medium, high and very high for classification ([15], [17]).The purpose of this paper [18] was to classify soil fertility in three levels using machine learning. This study was made in the hilly areas of Uttarakhand (India).This study [19] also used an unbiased predictor to predict soil organic carbon. Soil fertility rating was measured with pH, micro and macro nutrients like N, Cu, Fe, K, P, OC and Zn by statistical technique called Bayesian network [20]. A wide range of statistical approaches [21] were able to analyze soil quality that is directly involved in good farming and crop production with good crop health. [14] focused to apply a combination of 20 classifiers to classify various soil nutrients and fertility indices. Machine Learning can be widely used in each area of research for classification and prediction. The nature of ANN is to make a complex structure with mathematical functions. Different models were used in this paper [22] for pattern recognition. ANNs is very complex structure having mathematical approaches. It can be used to simulate human learning and pattern recognition). In India, The information reports about soil fertility are produced at district level or block level. These block wise reports help to soil analyst experts for making decisions about input fertilizers. These reports also help to make guidelines of distribution and utilization of fertilizers. In the current study we do classification of soil fertility in the context of several nutrients. This practice of such classification is useful to achieve an index report which is depends on the area wise soil fertility of Bhimtal. It would be helpful for making recommendations about the fertilizer in decision making system. A comparative study among the villages for soil fertility levels can be done by using this report. Therefore soil fertility classification for Organic Carbon (OC), Phosphorus (P) Potassium (K), Manganese (Mn) and Boron (B) is the main objective of our study. Once above mentioned nutrients are get classified then predictions

come in the next task for soil features using machine learning approach. These predictions always help to reduce the unnecessary expenditure on fertilize. It also saves the time of domain experts in soil health analytics.

2. SOIL FERTILITY PREDICTIONS FOR DIFFERENT VILLAGES IN BHIMTAL

We have focused our study on Bhimtal block in Uttarakhand. Bhimtal is located in Nainital district of Uttarakhand. It is a rural part of Uttarakhand. Most of the land in Bhimtal block is agricultural and surrounded with hills. The soil lands of Bhimtal block are good of cultivation and suitable for agriculture. Soil fertility is a primary aspect of soil productivity. Due to large number of exercises by inadequate or excessive use of fertilizers on agricultural soil decrease the level of fertility in the land. Soil fertility problems can be resolved by understanding the properties of soil lands. In this study we are also interested to predict the levels of soil parameters of various villages of Bhimtal block using Machine Learning techniques. Prediction of soil parameter levels also aids to avoid unnecessary expenditure on fertilizers. This prediction can save time of soil analysts for analyzing the soil health and environmental conditions. We have used deep neural network to classify fertility indices for five nutrients with two activation functions. We quantified soil nutrients values on the basis of Indian standard rating [23] in three categories Low, Medium and High. Inspired by this paper [15] we have designed an Artificial Neural Networks with two activation functions to determine fertility index for the selected five nutrients including (OC, P, K, Mn, and B). We have used machine learning methods to determine the soil fertility index for a soil nutrient given by Rammoorthy and Bajaj [24] in our data set. The formula of finding the Fertility Index is written as:

$$(2.1) \quad FI = \frac{N_L + 2N_M + 3N_H}{N_T}$$

The number of cultivated lands quantified as Low, Medium and high level group of nutrients for any specific village is denoted by N_L , N_M and N_H respectively. The soil fertility index for these nutrients can be calculated by above formula (2.1). We can divide this procedure to calculate FI in two steps as follows:

Step 1: We have taken the soil samples from the villages of the Bhimtal block. Each village has minimum one cultivation land. So in the first step each cultivation land of each village will be evaluated. This evaluation can be done on the basis of its fertility for the nutrients of both types i.e. micro and macro. The numeric values of each input nutrients for each cultivation land was represented in the ordinal values (Low, Medium and High) given by Agricultural department. The three tier ratings of nutrients are given by [23] and defined in Table 1.

Step 2: In this step an effort has been made to understand the intricacies of fertility levels of cropped land. In this regard N_L , N_M and N_H denoting low fertility, medium fertility and high fertility of cultivation lands

and for the purpose of total number of cultivated lands in each village $N_T = N_L + N_M + N_H$ was computed. Based on formula above in the equation (2.1) the fertility index (FI) was find out for each village taken into the consideration of the study. The calculated FI is an average of number of lands with the low, medium and high values [15]. We achieved fertility indices for the nutrients between 1 and 3 that means near to 1 fertility index refers to low fertility, near to 2 fertility index refers medium fertility and near to 3 refers high fertility for a individual nutrient of a village. Thus by using this procedure we can determine the fertility index for the nutrients with their rank. We have used the standards intervals for the various soil nutrients given by Govt. of India. These standard intervals are mentioned in Table 1 given below. The intervals for OC (%), P (kg/ha), K (Kg/ha) are given Table 1. The intervals for micro nutrients Mn and B (in parts per million, PPM) ([15], [23]) are given in Table 2.

Organic Carbon(OC)%	Primary Soil Nutrients		Ratings
	P(Kg/ha)	K(Kg/ha)	
0.21-0.40	7.1-14	101-150	Low
0.41-0.60	14.1-21	151-200	Medium
0.61-0.80	21.1-28	201-250	High

TABLE 1. Availability of fundamental nutrients (P, K and OC) in three tier rating [23]

Mn(ppm)	B(ppm)	Ratings
<2.5	0.25	Low
2.6-4.5	0.6-0.75	Medium
>4.6	0.76-1.0	High

TABLE 2. Three Tier rating of Manganese (Mn) and Boron (B) [15], [23]

Table 1 displays the three levels of ratings of Organic Carbon (OC) and the primary nutrients of soil P and K. In table 2 the three levels of ratings of micro nutrients Mn and B available in soil have shown. These ratings are determined and provided by Agriculture planning of Indian government.

3. MATERIAL AND METHODS

It would be expedient and necessary here to give brief discussion of applied appropriate methodology, tools and techniques. We have also discussed the study area, soil data set we have used in Bhimtal block. This paper covers the prediction fertility indices of selected available soil nutrients like organic carbon (OC), Phosphorus (P), Potassium (K) and micro nutrients Manganese (Mn), Boron (B). In our study the output of each nutrient should be predicted by machine learning.

3.1. Study Layout. The Himalayan state of Uttarakhand is located in northern part of the country and having the international edges with Tibet, China and Nepal ([25], [26]). Total area of the state of Uttarakhand is about $53,483 \text{ km}^2$ and It has an area of $53,483 \text{ km}^2$ and lies between latitude $28^\circ 43'$ and $31^\circ 28'$ N and longitude $77^\circ 34'$ and $81^\circ 03'$ E ([25], [26], [27]). Nainital district of Uttarakhand is situated between $28^\circ 8'$ and $29^\circ 6'$ North latitude and between $78^\circ 8'$ and $80^\circ 14'$ East longitude and covering an area of about 4251 square kilometers [26].

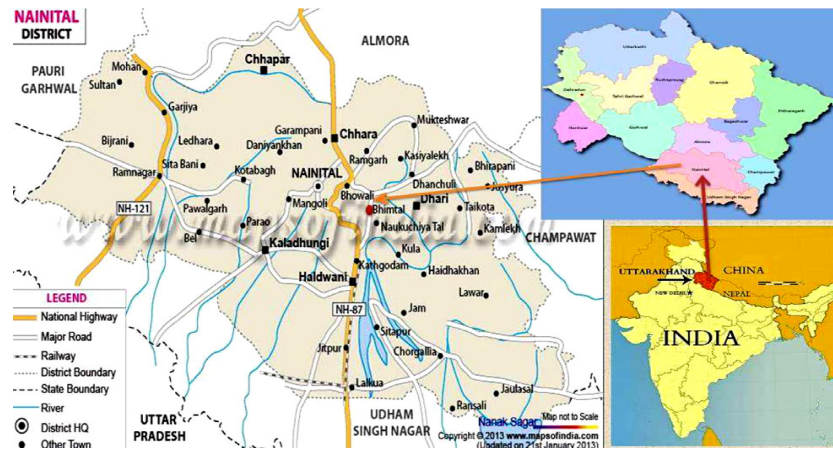


FIGURE 1. Geographic map of the Nainital district, dark dot indicate the Bhimtal our study area [26]

The Nainital district can be divided in two regions on the basis of its geographical conditions. These two areas are hill and bhabhar. The maximum area of Nainital districts are Hilly areas. In terms of agriculture activities and cultivation of crops in the Bhimtal region it is indispensable to highlight here that the climate of this region vary between alpine to sub-tropical geographical zone. Due to variability among villages in Bhimtal Block further the climate varies from hot to very cold [26].

A big number of populations of Bhimtal have migrated to the different cities of India from villages for their livelihood. But still a major source of income is agriculture for Bhimtal people's livelihood.

3.2. About Soil Data Set. The data samples we have applied in our study were gathered from the soil testing laboratory, Bhimtal Uttarakhand during the year 2018-19. The soil data is analyzed for 12 parameters including soil reaction pH, Organic carbon (OC), Electrical Conductivity (EC) primary nutrients N, P, K secondary nutrients S and micro nutrients Zn, Fe, Cu, Mn and B [15]. The input parameter EC is measured in milli-siemens per centimeter (ms/cm), OC is measured in %, and N, P and K are measured in Kg/ha [15]. All micro nutrients, Cu, Fe, Mn, Zn and B are measured in ppm (parts per million) [15].

Soil Data set contains dependents variables and independents variables. We get dependent variables by equation (2.1) for each cultivation land of each village. Further we have set the notation for soil fertility

indices for the dependent variables as, P-F, OC-F, Mn-F, K-F, and B-F for Phosphorus, Organic Carbon, Manganese, Potassium and Boron respectively. Both types of variables are present in our study so this study is known as multivariate study. We have obtained the fertility index of dependent variable by using the equation given in (2.1).

Total 946 samples are there in our data set each sample is a numeric vector with 12 input parameters for a particular area in block.

3.3. Village wise Classification of soil nutrients in terms of fertility indices. The unnecessary experiments in the lands during the cultivation may cause huge losses in soil in terms of productivity and fertility [15]. The experiment such as liming can be used to improve the acidity of soil. But this exercise is not given enough consideration in Uttarakhand because the soil is induced by all these factors. So it is concluded as the deficiency or excess of any soil nutrients may cause complications in soil.

It becomes a big challenge for the farmers and soil scientists to balance the nutrients level according to crop so that crop production can be grow with good crop health. These types of challenges can be dealt by machine learning to classify soil nutrients of each village for determining soil fertility. These fertility indices are useful for preserving the important soil nutrients. Fertility indices are also very important to determine the deficiency and excess of nutrients in soil. Soil samples were collected from Soil Testing Laboratory in Bhimtal. We have collected 946 samples from 85 villages in which 722 samples were with high level of OC.

Each nutrient has a specific property that is required for the better crop production. For photosynthesis process plants absorb P from soil. The amount of P is low in our study area. For plant development and better metabolism K plays a vital role. K is in medium level for our collected samples. We found 874 samples in which S is sufficient. So generally we can say that S is sufficient for the soils of Bhimtal block in Uttarakhand. A very small quantity of Zn is sufficient for a plant. It is useful for the crop's growth by producing proteins [28].

Plants also require Cu in small quantity and it is an ingredient of necessary enzymes. The physiology of plants depends on the amount of B. In this study we are working on five fundamental nutrients. The fertility index level of all primary, micro and macro nutrients are defined and describes by agricultural planning of Indian Government [23]. In this study we use the standard value of OC, P, K, Mn and B provided by agricultural planning of government of India.

The fertility level of nutrients can be quantified as in three levels of Low fertility, Medium Fertility and High fertility, discussed earlier, by using the threshold values of respective nutrients given by agricultural planning. The fertility index values near to 1 is considered as low level fertility, the fertility index values near to 2 is considered as medium level fertility, the fertility index values near to 3 is considered as high level fertility. The range values of fertility of soil for area wise is formed in the given Table 3.

Fertility Index	Fertility Index
<1.67 i.e. near to 1	Low
1.67-2.3 i.e. near to 2	Medium
>2.33 i.e. near to 3	High

TABLE 3. Threshold values for fertility indices [23]

A method [24] is used to compute the village wise soil fertility index. The formula is defined above in equation (2.1). The Soil samples were taken from number of cultivation fields of each village of Bhimtal block. For any particular village the Fertility Index (FI) value is same for all cultivation fields. The target values are fully distinct from the values of village wise fertility index of these values.

For our classification problem, the soil fertility indices inputs which are used in this paper are also recorded in Table 4.

Soil Fertility Indices Classification Problem in villages of Bhimtal	Levels/Class	No of Cultivation lands in per Level
OC-F	Low	125
	Medium	99
	High	722
P-F	Low	557
	Medium	296
	High	93
K-F	Low	196
	Medium	599
	High	313
Mn-F	Low	330
	Medium	103
	High	513
B-F	Low	426
	Medium	104
	High	416

TABLE 4. Summarization of Fertility Index and Cultivation Lands in Various Class Levels

3.4. Expected Frame work of soil classification. In this section a framework is designed to implement a model for the above classification problems mentioned in Table 4. The data was 0.9 split factor i.e. 90% data for training and validation and rest of data i.e. 10% for testing purpose. A cross validation, tenfold, method is used for training and validation of the data. For training and validation, a tenfold cross validation method is used for training and validating the data. The division of training data was done at 9:1 ratio i.e. 90% of data for training and 10% of trained data for validation. This division was set same for each fold. In this study ANN classifiers are used to train the model by using the different activation functions and hidden nodes in ANN architecture. After completion of training process, the classifiers are verified for the validation set. Then the best parameters are chosen from training set and then it is used in the test set of the data. The output of the corresponding classifier is considered as the final test result and used for analysis. The strategy for the classification is designed in Figure 2.

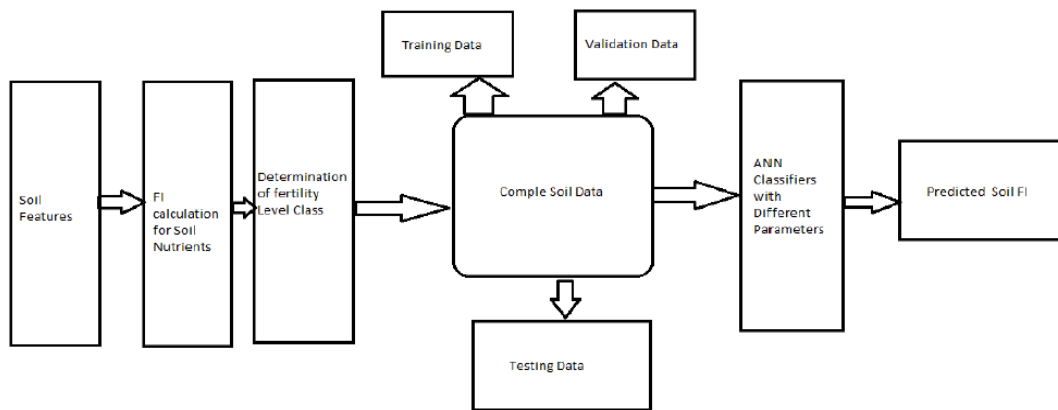


FIGURE 2. Expected Prototype for Soil Features Classification and Prediction

The different sets of parameters are used to train each ANN classifier on the training data. The parameters are training functions, input, hidden and output layers. The ANN classifiers are used on the validation data after the training and obtained best parameters are used for testing the data [15]. The final output of the classifier is the achieved best accuracy of test data. Two different Meta parameters are used for the ANN deep networks and number of epoch is a Meta parameter for optimization. Number of epoch parameter will remain fixed for all the classifiers for nutrients classification and prediction in soil. Activation functions are other parameters used in this study are hyperbolic tangent and rectified linear unit. Once the whole data is broken down in training and test, the classifier is used in the training data having the activation functions for the classification process. The test data is used to make predictions and calculate the accuracy of classifier. Different used models can be compared by obtaining the accuracy of each model and the activation function of best accuracy is treated as optimal parameter of the model or classifier.

4. RESULT AND DISCUSSION

Soil fertility indices for five nutrients and their classification in three different classes are described in Table 3. This region has low and high level of OC-F. The soil consists of low and medium level of P-F and medium and high level of K-F. The soil of Bhimtal region contains the Mn-F and B-F with low and high level.

Pursuant to geographical pattern of Bhimtal Block most of the villages are located in hilly region and there is always the possibilities of soil erosion and leaching of soil nutrients. The soil data we have used in our study is preprocessed by calculating the standard deviation and mean [16] of each input feature. The result of these preprocessing is plotted in Figure 3.

	pH	EC	OC	N	P	K	S	Zn	Fe	Cu	Mn	B
count	947.000000	947.000000	947.000000	947.000000	947.000000	947.000000	947.000000	947.000000	947.000000	947.000000	947.000000	947.000000
mean	7.220063	0.116325	1.067648	0.206304	13.980697	257.631447	13.132396	1.448414	10.380982	0.539317	5.034831	0.514761
std	0.494968	0.184005	2.943864	5.072913	21.691627	187.185749	5.470020	1.023916	12.854929	0.964565	6.829647	0.868599
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	7.000000	0.000000	0.800000	0.000000	4.500000	156.500000	10.760000	0.695000	4.747500	0.231000	0.891000	0.000000
50%	7.300000	0.000000	1.070000	0.000000	9.000000	216.000000	12.410000	1.104000	6.960000	0.345000	4.742000	0.000000
75%	7.500000	0.300000	1.210000	0.000000	18.000000	284.000000	14.805000	2.060000	9.630000	0.550000	7.636000	1.035000
max	8.400000	1.200000	91.000000	156.000000	400.000000	2514.000000	140.700000	6.083000	77.290000	10.900000	176.000000	10.000000

FIGURE 3. Preprocessing of Data

The positive or negative relation between two variables can be defined by correlation analysis. The correlation between each pair of attributes is represented by correlation matrix. A correlation matrix for each pair of soil nutrients is shown in Fig 4. Correlation matrix represents how two variables are highly correlated with each other. Some machine learning models give poor performance if highly correlated variables are present in the data. So it is useful to understand the correlation matrix. We can see from the Figure 4 the bottom left of matrix is same as the top right matrix. So this correlation matrix is called as symmetrical matrix.

A scatter plot is another approach to visualize data preprocessing. It is used to show the relationship between two features as dots in two dimensions. Basically scatter plots are useful to exhibiting a structured relationship between the features. So we can outline the relationship with a line between the two features. There may be correlation between these features with a structured relationship and these features are good for the elimination from the data. Scatter matrix is symmetrical same as the correlation matrix. If we want to show the pair wise relationship between the features from the different perception then this scatter matrix is useful. The scatter plot for each pair of attributes of soil data can be understood in Figure 5.

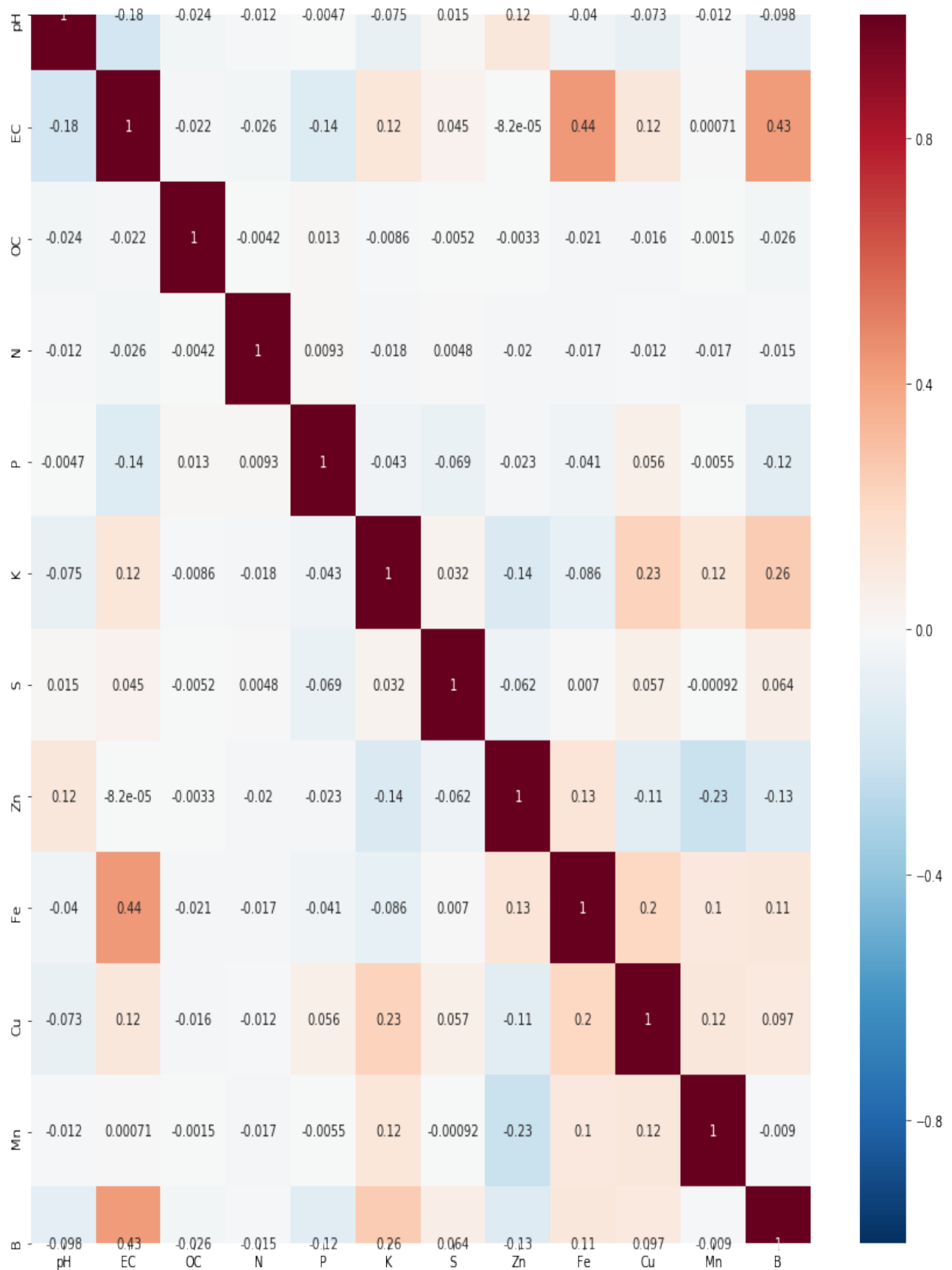


FIGURE 4. Correlation Matrix

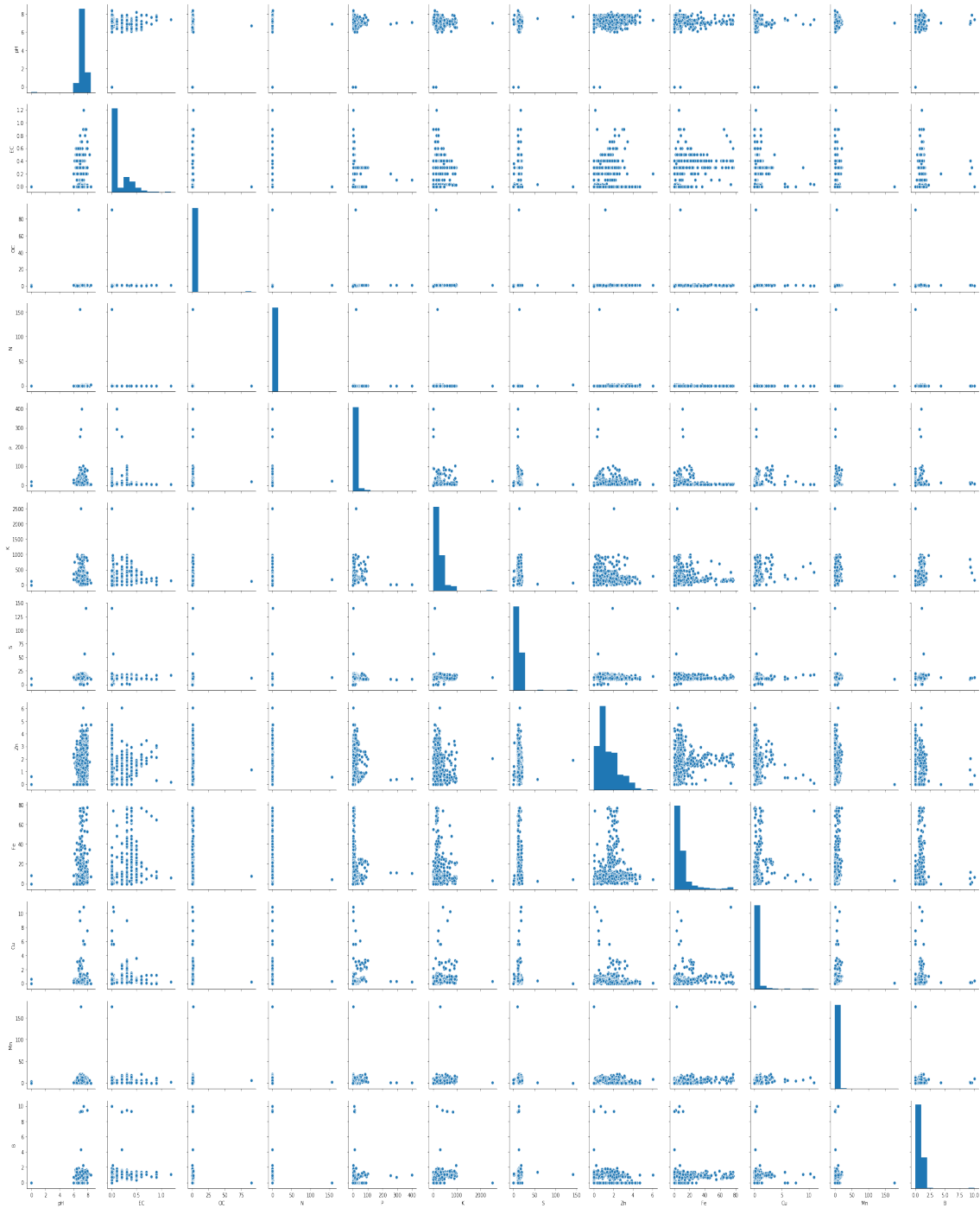
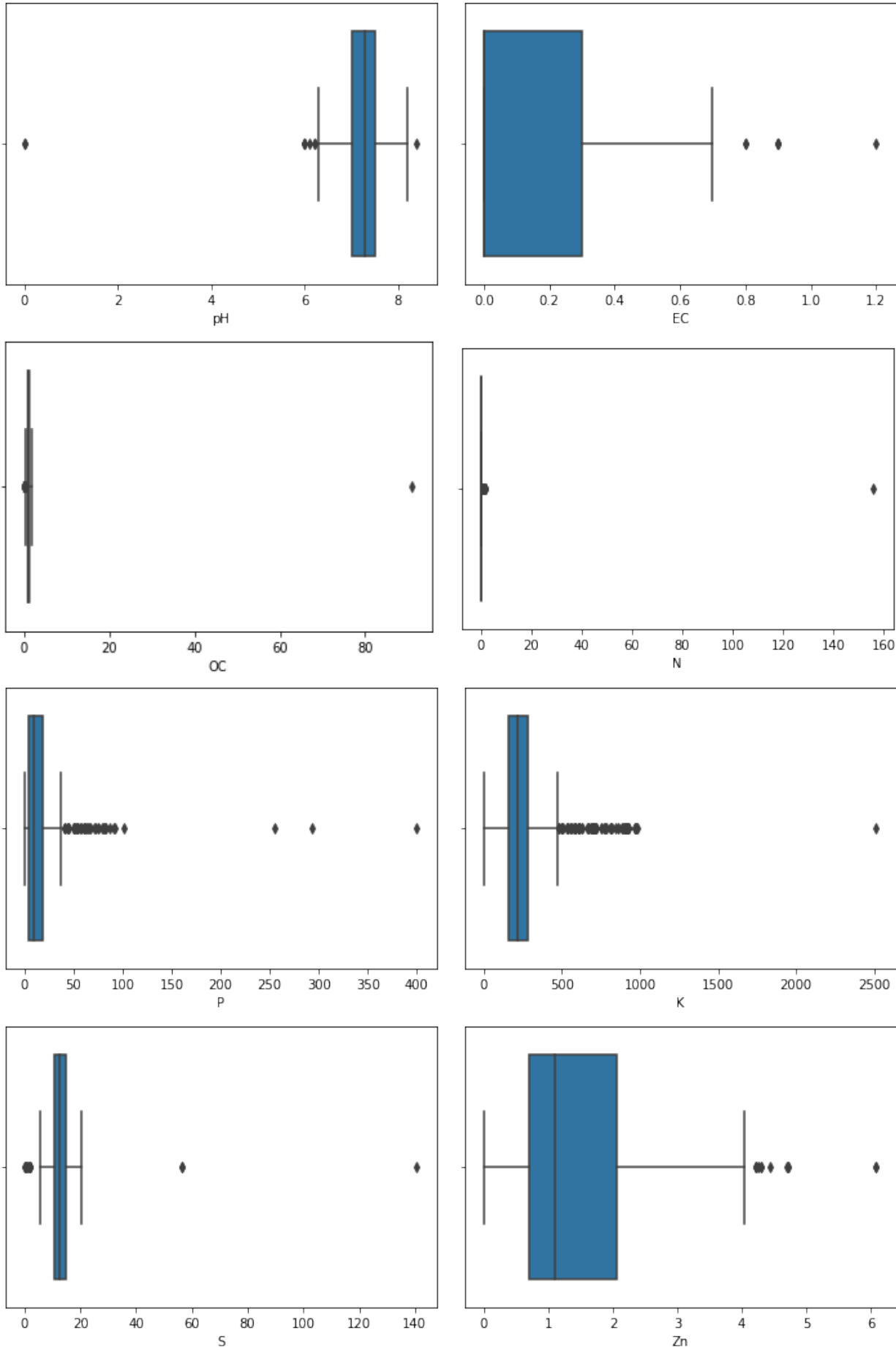


FIGURE 5. Scatter Plot of the Data



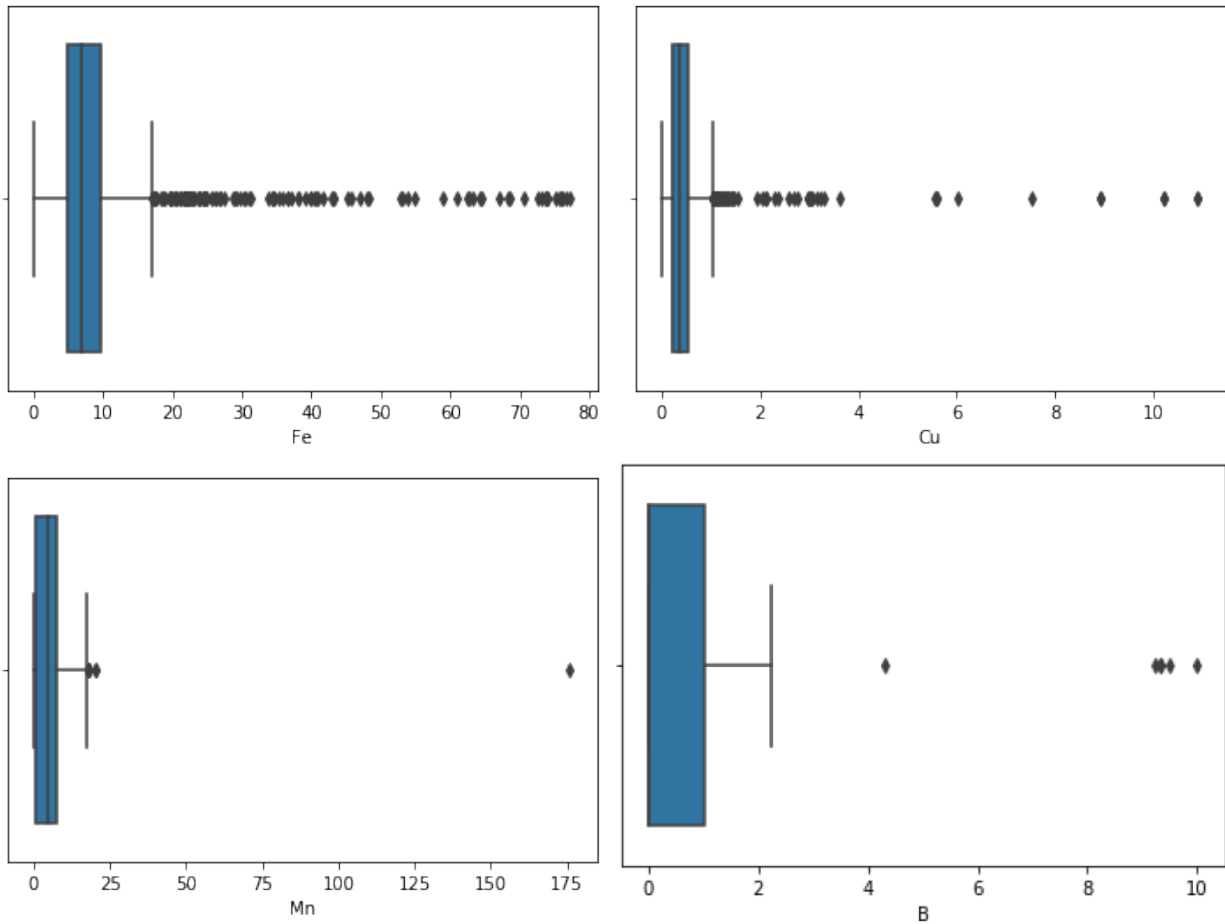


FIGURE 6. Box Plot of Input Feature and Identification of Outliers

Box plots of all the twelve input features can be understood by Fig 6. The input data is describing in 25% and 75% quartiles. The median of the each input feature is represented by black line inside the box. The extremely less and extremely high values which are far from the box plot are considered outliers. In our study we used Inter quartile range method to remove outliers' from the data. For each classifier these input features would be same but output of the prediction of each classifier would be different.

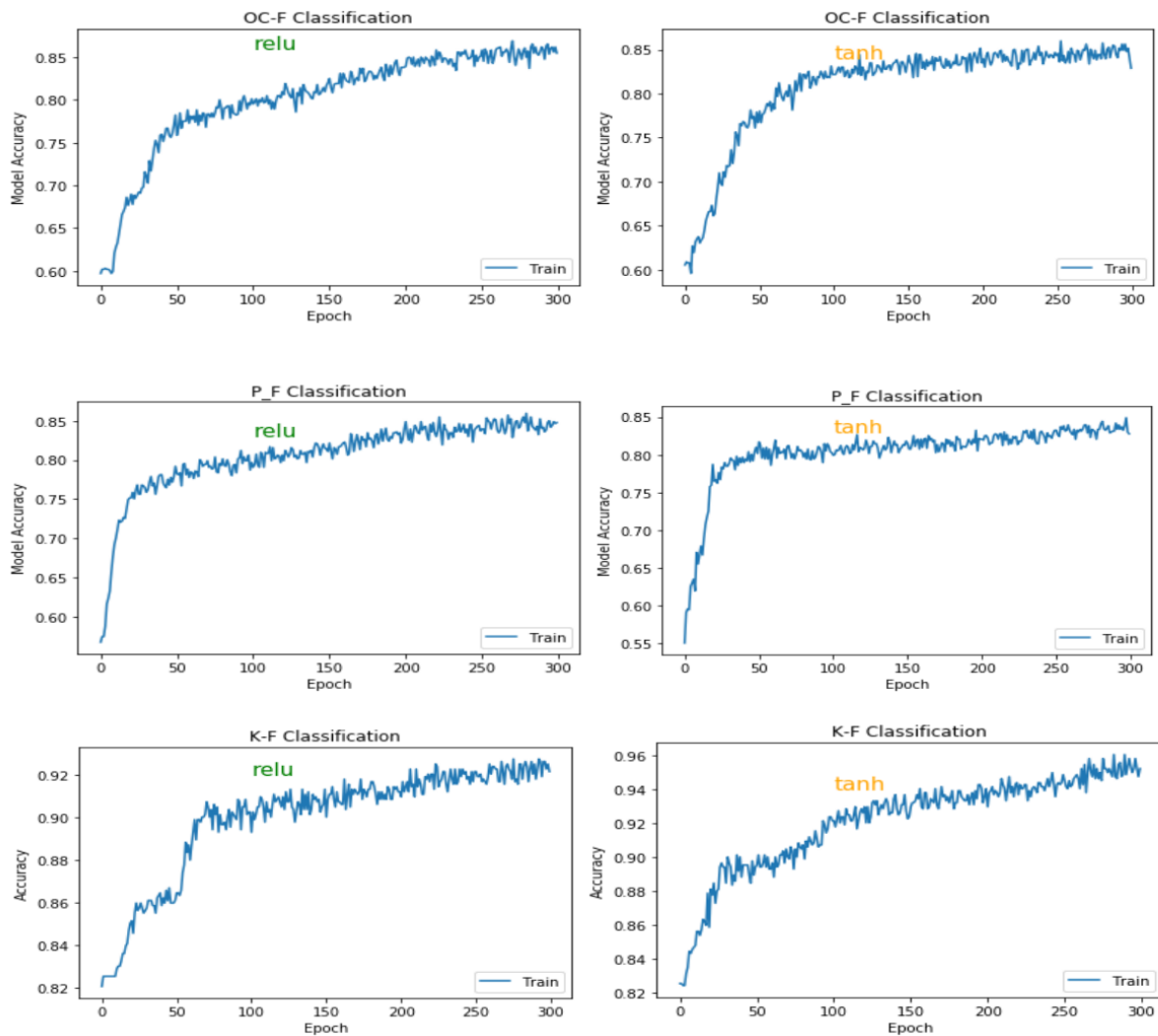
In the current study we use the fast learning algorithms of deep learning. We have used Keras library along with the relu, tanh activation functions as classifiers for the classification of soil in python. We have used deep neural network with two activation function i.e. Rectified Linear Unit (Relu) and Hyperbolic tangent (tanh) to analyze and classify our problem. The performance of these two activations functions is also compare for all the village soil fertility indices. The performance of the classifiers can be measured by the accuracy of the classifiers. It computes by the correct count of the predictions. The accuracy of the classifier is calculated by the given formula. Accuracy is a fraction of the exact predictions over K samples

and defined as

$$(4.1) \quad accuracy(Y, y) = \frac{1}{k_{samples}} \sum_{i=0}^{k_{samples}-1} 1(y_i = Y_i)$$

Where predicted value of i^{th} sample is denoted by y_i , and actual true value is denoted by Y_i and $1(x)$ is an indicator function [29].

The results of classification problems for five fertility indices are plotted in Fig 7 below. The highest accuracy of the classifier is indicated as the peak value in the graph in Fig 7 and cross validated accuracy obtained by classifiers is also described. This accuracy is achieved for the five classification problems. From the plotted graph we can see that the best performance of soil fertility classification is achieved by rectified linear unit (relu) function for four classification problems and one classification problem is handled by hyperbolic tangent (tanh) function with best accuracy. The accuracy achieved by the activation function can be understood by Table 5.



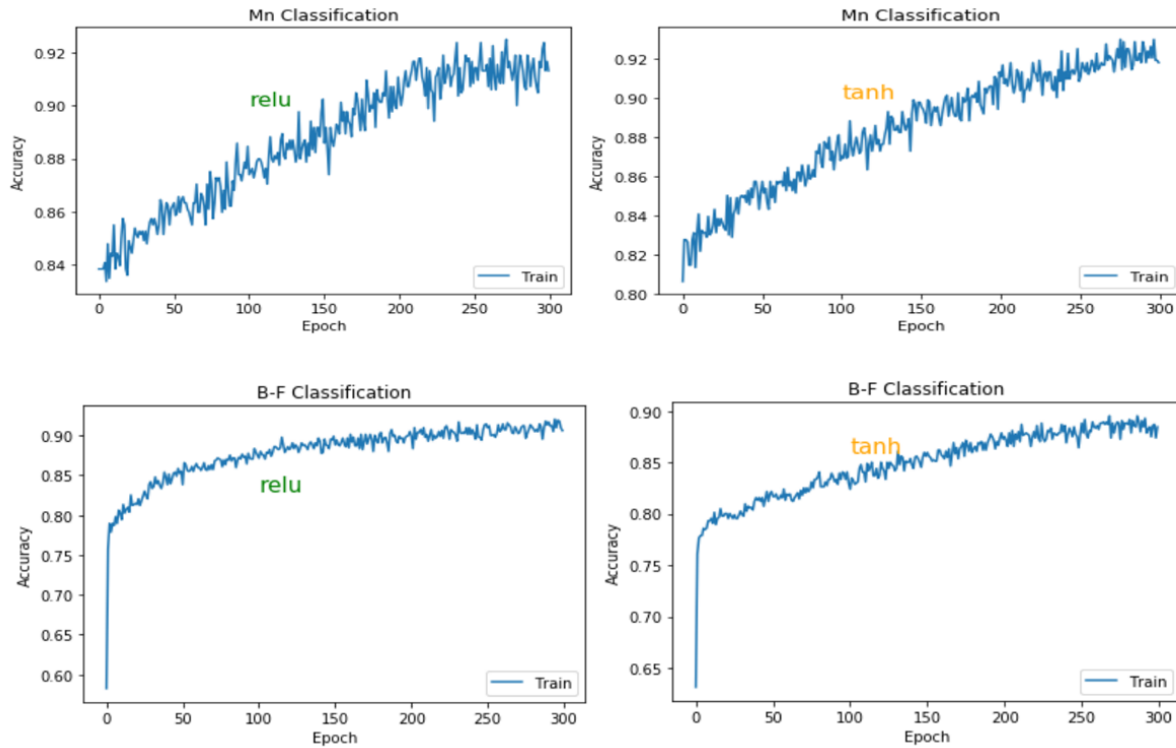


FIGURE 7. Accuracy Comparison

Classifier	Soil Nutrients Accuracy				
	OC-F	P-F	K-F	Mn-F	B-F
Relu	91.30	84.74	93.15	91.32	90.00
Tanh	80.12	81.90	95.52	91.80	88.47

TABLE 5. Achieved ANN Classifier accuracy for soil fertility classification problem. The best accuracy for each classification problem is highlighted in bold.

Now It can be understood that relu obtained the better accuracy for fertility classification of OC, P, Mn and B while tanh obtained better accuracy for fertility classification of K.

5. CONCLUSION

Poor literacy level, lack of awareness and inadequate knowledge with regard to fertilizers, pesticides and insecticides among farmers is leading to quality deterioration in terms of soil fertility. The composition of nutrients in soil may be ineffective if higher quantities of agri-inputs are supplied to soil. In spite of enhancing the productivity of crops, it is expedient and necessary to conserve soil and other natural resources.

In order to get meaningful, appropriate, reliable, valid and authentic results we employed disproportionate stratified random sampling distribution technique. It enables us to perform proper analysis of factors required for any crop and soil so that genuine and meaningful usage and practices regarding improved soil fertility and enhanced crop production can be understood properly.

Pursuant to geographical pattern of Bhimtal Block most of the villages are located in hilly region and there is always the possibilities of soil erosion and leaching of soil nutrients.

The main objective of the study is to categorize village wise soil fertility indices for the various nutrients of Bhimtal, a promising hilly area of Uttarakhand. This study will help the stakeholders in decision-making through improved awareness regarding soil quality and crop production. The soil quality depends on pH, EC, primary, micro and macro nutrients for a particular crop. Therefore, classification of soil nutrients helps to save the time of soil analysis experts and farmers. In this study we have made an attempt of machine learning for classifications of soil problems by an efficient manner in the area of Bhimtal Block. This study permits us to apply Deep ANN networks for classification of fertility level of soil along with two activation functions. Further results observed and measured are depicted in three levels as: Low Medium High. In addition to this another attempt has been made to analyze the accuracy of two activation functions (relu and tanh) of ANN classifier. Results are self-explanatory. It is expedient and necessary to mention here out of five classified soil nutrients parameters relu activation function has shown better performance among four classified soil nutrients parameters while tanh proved better performance in only one classified soil nutrients parameters. Among all selected 85 villages the accuracy level of OC-F is found to be more than 91% in fertility index level, accuracy level of P-F is found to be more than 84%, accuracy level of K-F is found to be more than 95%, accuracy level of Mn-F is found to be more than 91% and accuracy level of B-F is found to be more than 90%. All above discussed nutrients are essential ingredients for any kind of crop cultivation especially cash crop. Based on the above analysis and discussion it can be generalized that farmers of all selected villages from Bhimtal Block can be benefited if proper developmental strategy regarding these nutrients are disseminated, conveyed in the prevalent dialects. It has been found that the relu activation function provides best performance to examine the level of accuracy of soil fertility in respect of the nutrients of OC, P, Mn and B while tanh is the best performer to examine the accuracy of soil of nutrient of B. The implementation of the classifier has been done in python by using the deep learning libraries such as keras and tensorflow. Both classifiers achieve good results in the classification of five soil problems. Consonance with statistical values relu function provides best results for OC-F, P-F, Mn-F and B-F classification with the prediction and estimation accuracy 91.30%, 84.74%, 91.32% and 90% respectively while tanh provides the best result for K-F with the prediction and estimation accuracy of 95.52%.

Conflicts of Interest: The author(s) declare that there are no conflicts of interest regarding the publication of this paper.

REFERENCES

- [1] National Research Council. *Alternative Agriculture*. National Academy Press. Washington, D.C. 1989.
- [2] P. Gruhn, F. Goletti, Yudelman M *Integrated nutrient management, soil fertility, and sustainable agriculture: current issues and future challenges*. International Food Policy Research Institute, Washington, D.C. 2000.
- [3] A. Mucherino, P. Papaorgji, P.M. Pardalos, A survey of data mining techniques applied to agriculture. *Oper. Res.* 9 (2009), 121–140.
- [4] K. Lamorski, Y. Pachepsky, C. Slawiński, R.T. Walczak, Using support vector machines to develop pedotransfer functions for water retention of soils in Poland. *Soil Sci. Soc. Amer. J.* 72 (2008), 1243–1247.
- [5] M. Kovačević, B. Bajat, B. Gajić, Soil type classification and estimation of soil properties using support vector machines, *Geoderma*. 154 (2009), 340–347.
- [6] Y. Liu, H. Wang, H. Zhang, and K. Liber, A comprehensive support vector machine-based classification model for soil quality assessment, *Soil Tillage Res.* 155 (2016), 19–26.
- [7] Y. Huang, Y. Lan, S.J. Thomson, A. Fang, W.C. Hoffmann, R.E. Lacey, Development of soft computing and applications in agricultural and biological engineering, *Comput. Electron. Agric.* 71 (2010), 107-127.
- [8] W. Wu, A.D. Li, X.H. He, et al. A comparison of support vector machines, artificial neural network and classification tree for identifying soil texture classes in southwest China. *Comput. Electron. Agric.* 144 (2018), 86–93
- [9] H. Li, W. Leng, Y. Zhou, F. Chen, Z. Xiu, D. Yang, Evaluation Models for Soil Nutrient Based on Support Vector Machine and Artificial Neural Networks, *Sci. World J.* 2014 (2014), 478569.
- [10] M.H. Yang, D.Y. Xu, S.C. Chen, et al. Evaluation of machine learning approaches to predict soil organic matter and pH using Vis-NIR spectra. *Sensors*. 19(2) (2019), 263.
- [11] M. Hosseini, S. Rajabi Agereh, Y. Khaledian, et al. Comparison of multiple statistical techniques to predict soil phosphorus. *Appl. Soil Ecol.* 114 (2017), 123–131.
- [12] K. Liakos, P. Busato, D. Moshou, S. Pearson, D. Bochtis, Machine Learning in Agriculture: A Review, *Sensors*. 18 (2018), 2674.
- [13] J.R. Romero, P.F. Roncallo, P.C. Akkiraju, et al. Using classification algorithms for predicting durum wheat yield in the province of Buenos Aires. *Comput. Electron. Agric.* 96 (2013), 173–179.
- [14] M. Sirsat, E. Cernadas, M. Fernández-Delgado, M. Khan, Classification of agricultural soil parameters in India, *Comput. Electron. Agric.* 135 (2017), 269–279.
- [15] M.S. Sirsat, E. Cernadas, M. Fernández-Delgado, S. Barro, Automatic prediction of village-wise soil fertility for several nutrients in India using a wide range of regression methods, *Comput. Electron. Agric.* 154 (2018), 120–133.
- [16] M.S. Suchithra, M.L. Pai, Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters, *Inform. Proc. Agric.* 7 (2020), 72–82.
- [17] X.E. Pantazi, D. Moshou, T. Alexandridis, R.L. Whetton, A.M. Mouazen, Wheat yield prediction using machine learning and advanced sensing techniques. *Comput. Electron. Agric.* 121 (2016), 57–65.
- [18] J. Pant, P. Pant, A. Bhatt, H. Pant, N. Pandey, Feature Selection towards Soil Classification in the context of Fertility classes using Machine Learning, *International Journal of Innovative Technology and Exploring Engineering*, 8 (12) (2019), 4000-4004.

- [19] C. Ritz, E. Putku, A. Astover, A practical two-step approach for mixed model-based kriging, with an application to the prediction of soil organic carbon concentration. *Eur. J. Soil Sci.* 66 (2015), 548–554.
- [20] H.-Y. Jia, J. Chen, H.-L. Yu, D.-Y. Liu, Soil fertility grading with Bayesian Network transfer learning, in: 2010 International Conference on Machine Learning and Cybernetics, IEEE, Qingdao, China, 2010: pp. 1159–1163.
- [21] D. Elavarasan, D.R. Vincent, V. Sharma, A.Y. Zomaya, K. Srinivasan, Forecasting yield by integrating agrarian factors and machine learning models: A survey. *Comput. Electron. Agric.* 155 (2018), 257–282.
- [22] T. Kim, Pattern Recognition Using Artificial Neural Network: A Review, in: S.K. Bandyopadhyay, W. Adi, T. Kim, Y. Xiao (Eds.), *Information Security and Assurance*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010: pp. 138–148.
- [23] Integrated Nutrient Management Division. *Methods manual: soil testing in India*. Department of Agriculture and Cooperation, Ministry of Agriculture, Government of India, New Delhi, pp 89–96. (2011).
- [24] B. Rammoorthy, J. Bajaj. Available N, P and K status of Indian soils. *Fertilizer News* 14 (1969), 24–26.
- [25] O.P. Tiwari, Y.S. Rana, R. Krishan, C.M. Sharma, B.S. Bhandari, Regeneration dynamics, population structure, and forest composition in some ridge forests of the Western Himalaya, India, *Forest Sci. Technol.* 14 (2018), 66–75.
- [26] K. Gariya, G.K. Dwivedi, V. Kumar, S.K. Tewari, Socio-economic Characteristics of Homegardens in Bhimtal block of Nainital District, Uttarakhand, India, *Int. J. Agric. Environ. Biotechnol.* 9 (6) (2016), 1001-1013.
- [27] J. Sharma, S. Gairola, R.D. Gaur, et al. Plants used for treatment of dysentery and diarrhoea by the Bhoja community of district Dehradun, Uttarakhand, India. *J Ethnopharmacol.* 150 (2013), 969–1006.
- [28] P. Arunachalam, P. Kannan, G. Prabukumar, M. Govindaraj, Zinc deficiency in Indian soils with special focus to enrich Zinc in peanut. *Afr. J. Agric. Res.* 8 (2013), 6681–6688.
- [29] M. Sokolova, N. Japkowicz, S. Szpakowicz. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In: *Australian conference on artificial intelligence*, vol 4304. LNCS, Germany, pp 1015–1021. (2006).