

Soil Nutrients Prediction and Optimal Fertilizer Recommendation for Sustainable Cultivation of Groundnut Crop using Enhanced-1DCNN DLM

Sivasankaran S¹, Dr. K. Jagan Mohan², Dr. G. Mohammed Nazer³

Research Scholar, Department of Computer Science and Engineering
Annamalai University, Chidambaram, Tamilnadu, India¹

Associate Professor, Department of Information Technology
Annamalai University, Chidambaram, Tamilnadu, India²

Professor and Principal, Department of Computer Science
RAAK Arts and Science College, Villupuram, Tamilnadu, India³

Abstract—Cultivation of crops and their parallel production yields hugely depend upon the fertility composition of the soil in which the crops are being cultivated. The prime fertility factors which contribute towards the health of the soil are the available soil nutrients. Varying climatic conditions and improper cultivation patterns have resulted in unpredictable growth and yield of the groundnut crops, one of the major cause for the fluctuation seen in groundnut pod growth patterns and production, is the differing soil nutrient compositions of the land which is under cultivation. The unnecessary usage of excessive artificial fertilizers to boost the soil strength, without properly diagnosing the exact nutritional need of the soil required for the conducive growth of the crop has led to the imbalanced distribution of the soil's major macro-nutrients constituents such as (Phosphorous (P), Potassium (K) and Nitrogen (N)). In this research article, we have made a detailed investigation for nutrient prediction mechanism of the soil nutrient datasets taken under investigation of a specific geographic location from one of the major groundnut cultivating districts (Villupuram) in the state of Tamil Nadu and have proposed a Soil nutrients prediction scheme and optimal fertilizer recommendation model for sustainable cultivation of groundnut crop using Enhanced-1DCNN DLM. This Investigation model utilizes the natural compact robust features of 1DCNN in classifying the major macro nutrients(N,P,K)on the basis of low, Medium and high values. Based on the generated heatmap results the correlation between certain macronutrients and their corresponding micronutrient presence is classified. This proposed model has been compared for its performance and error measures with existing SVM, Naïve Bayes and ANN models and has proved to be outperforming all the compared baseline models by preserving the original data distribution with an overall accuracy of 99.78%.

Keywords—Soil nutrients; *Enhc-ID-CNN DLM*; nutrient classification; fertilizer recommendation

I. INTRODUCTION

Farming activities across the globe, has started taking a different dimension in its approach serving to the changing socio-economical needs. The impact of technology has already started to make inroads in the agriculture sector, several nations have sensed it and are slowly in the process of

adapting precision based agricultural activities. India being one of the world's largest agriculture-based country, the scope and possibility of adapting precision-based agriculture has slowly gained importance. Groundnut is one of the most predominant oil seed, which has been cultivated in our country. This nutritious nut has been cultivated across the year In India, in three seasons, namely, the monsoon or rainy season which is called as Kharif, the winter season which is called as Rabi and the Summer. In India, one of the most important groundnut-growing states is Tamil Nadu, where groundnut is grown in five seasons: Adipattam (June-July), Karthikaipattam (October-November), Margazhipattam (December-January), Masipattam (February-March), and Chithiraipattam (March-April) (April-May). The kharif season accounts for nearly 80% of the country's total groundnut production [25][26][27][28]. As of 2018-19, the area under groundnut cultivation in Tamil Nadu was around 3.38 lakh hectares. For the best results, groundnut cultivation requires sandy loam or clay loam soil with good drainage. The soil should be deep, the pH should be around 5.5 to 7, and the fertility index should be high. The heavy soil was found to be unsuitable for cultivation due to harvesting difficulties. Because these crops are salt-sensitive, the soil should not be salty. Groundnut crop soil should not contain any rocks or clay, as this will reduce the crop's yield during harvest. For healthy germination and growth, the temperature in the cultivation area should be around 27-30 degrees Celsius. The ideal annual rainfall for crops is between 450 and 1250 millimeters. Groundnut farming is not suitable for high altitudes, cold, and frosty climates. The cultivation of groundnuts benefits from a consistent warm climate. The cultivation and production of the groundnut have been affected by many factors, one of the important issues which leads the cultivation land to be less fertile is by the usage of excessive number of artificial fertilizers on to the soil, without properly predicting the exact soil nutrient deficiency, such cultivation patterns make the soil infertile over a period of time, by thus there is gradual decline seen in the production rates of the groundnuts. Growth of the groundnut crops and higher groundnut pod formations hugely depend upon the underlying soil nutrient composition on which the crop has

been cultivated. There are many macro and micro nutrients which contributes towards the productive health of the crops. The major set of macro nutrients which include (Nitrogen, Phosphorous and Potassium) NPK, defines a important role in defining the productive nutrient composition of farm in which the groundnut is being cultivated. In addition to this basic macro nutrients, calcium, Sulphur and few other micro nutrients such as iron, zinc and manganese also contributes towards production of healthy groundnut pods, which rises the overall production yield of the groundnut crop [28][29][30]. Impact of artificial intelligence by the means of Machine learning and deep learning models have started creating a unique dimension in addressing complex problems through various quantifiable solutions across diversifying domains. In this research article, we have proposed a deep learning model (Enhanced 1DCNN DLM), for predicting the various soil nutrients constituents required for the sustainable cultivation of the groundnut crop. Mathematical models developed to predict soil nutrients composition must, in general measure possible physical quantities of the environment and provides formulas to describe the actual relationships between these soil-related parameters. Because of advances in computer modelling, empirical modelling methods have emerged as a dominant development model capable of extracting contextual physical quantities related to soil nutrient composition [19][20][31][34]. This empirical model is frequently used to choose a good model, correct physical quantities using the chosen model, and validate it to see how accurate it is at predicting soil fertility. These models must take into account relevant data, such as predefined input parameters and required output parameters. The input parameters, in particular, are chosen empirically with the goal of maintaining a minimum degree of correlation between them. To build predictive models, the majority of the major soil nutrient prediction schemes that have contributed to the literature have primarily used neural networks, linear regression, machine learning, and empirical formulas [2][3][5][10][18]. Soil nutrient prediction models based on deep learning have been found to be more robust and reliable in terms of prediction accuracy at this time. There are considerable amount of research directions defined through various data mining and machine learning methodologies, towards soil nutrient based crop prediction, soil fertility prediction and for soil fertilizer recommendation [24][23][22]. Crop based Soil nutrient prediction schemes towards identifying the exact nutrient composition of the land selected for agriculture and precise fertilizer recommendations based on the prediction made seems to be a novel approach for implementation, and through deep learning model the prediction the performance metrics will be more and error measures will be comparatively less when compared with existing machine learning models [6][7][11][12].

II. RELATED WORK

Soil states a pivotal role towards determining yield of the crops, in recent years there where many data mining and machine learning-based investigations made on the basis of the soil nutrients availability and its proposed fertilizer recommendation schemes. These variety of machine learning based schemes have opened up a wide scope towards reaching

a more optimal soil nutrient prediction scheme more specific to particular crop cultivation patterns. Various base articles, pertaining to our research problem have been reviewed comprehensively in this context [13][14][15][16][17].

For classifications, Nikam et al.[1] defined a model involving J48/C4.5, knn, ID3, Artificial Neural Network, Support Vector Machine, and Naive Bayes. These classification methods are divided into three groups: statistical, machine learning, and neural networks. [1].

Three algorithms were used by J Gholap et al.[2] to define their work. J48(C4.5) and Jrip algorithms were used with 1988 soil instances in the J48: It is a very simple classifier that generates a decision tree with a 91.90 percent accuracy. The author also suggested that a future goal would be to develop a recommendation system that would suggest appropriate fertilizer based on the soil test sample and cropping pattern.

Dr. K. Arunesh et al. [9] investigated and experimented in 203 soil instances with 6 soil attributes from Virudhunagar District, Tamilnadu, India, and found that the Nave Bayes machine learning classification algorithm outperforms J48, random tree, JRip, and OneR,

Ramesh et al. [21] proposed a system that uses Naive Bayes, Bayes Net, Naive Bayes Updatable, J48, and Random Forest as classification algorithms. They used a dataset which comprised 1500 instances of Soil samples obtained from Department of Agricultural. In the classification of soil nutrients, J48 calculated 92.3 percent accuracy, while the Nave Bayes algorithm calculated 100 percent accuracy.

Using two algorithms, Nave Bayes and J48, Chiranjeevi M N et al. [5] proposed a system for analysing soil condition and nutrients which includes Potassium, pH, Nitrogen, EC, Phosphorus, OC, Sulphur, Iron, Zinc, Magnesium, Boron, and Copper at Belagavi Department of Agriculture in Belagavi. The Naive Bayes algorithm produced a better result than the J48 algorithm, correctly classifying the determined number of instances of the soil sample.

Naive Bayes Classifier had been applied to Tirupati, Andhra Pradesh soil, according to Bhargavi et al [8]. The soil data instances were all classified into different sand categories, such as loamy sand, clay, loam, sand, sandy loam, sandy clay loam, and clay loam.

Puno, J. C., et al. [18] proposed and developed a fully functional system using IP (IP enhancement, IP segmentation, and feature extraction) in MATLAB software. All 7 nutrient attributes are classified as L, M, H, S, D (Low, Medium, High, Sufficient, Deficient) values.

After experimenting with model 1 and model 2 for soil moisture estimation, Ahmad, S et al. [4] proposed a model based on five-year data with only one attribute for classification considered, namely VIC moisture. The author concluded that SVM model outperforms ANN and MLR models.

Juhi Reshma S R et al. [23] used Neural Networks to propose a recommendation system for predicting the number of fertilisers needed for a specific banana crop, as well as regression methods for upcoming plantations. Nitrogen (N),

phosphorus (p), and potassium (k) are the three most important soil nutrients for crop growth. By default, soil contains a specific amount of NPK, which varies by location. Each crop has its own set of requirements. A model is constructed in this paper to recommend the number of fertilisers required for the banana crop.

A. Extracts Inferred from the Literature

The limitations of the available soil nutrient prediction models considered towards the literature study across the recent years is drafted below.

1) Most of the available soil nutrient predicts models have been approached by data mining and machine learning algorithms, which devised the approach towards nutrient prediction, but there exist limitations in terms of learning accuracy.

2) Mostly the available machine learning schemes have not defined well defined a sustained prediction accuracy to fit the need, opening a wide scope for improvement.

3) Though there is a considerable amount of research contributions done towards predicting the soil nutrients and fertilizer recommendation, there is only a limited class of investigation done for crop specific soil nutrient prediction and fertilizer recommendation.

4) The comparison of error measures and performance metrics reached by the available ML defined soil nutrient prediction schemes extend a clear scope for improvement if learning can be further be deeply enhanced.

5) On the basis of the above drafted limitations, it has been proposed to propose an Enhanced -1D Convolutional Neural Network based Deep Learning Model (Enhanced-1DCNN DLM) to facilitate accurate estimation of soil nutrient prediction and fertilizer recommendation for the sustainable cultivation of groundnut crops.

B. Proposed Work and its Scope of Contributions

The major aids of the proposed Enhanced-1DCNN based Deep Learning Model is listed as follows:

1) The proposed model works towards the soil nutrients composition prediction and its necessary fertilizer recommendations, which aids towards the better cultivation of groundnut crops in the specific geographical location present in the Villupuram district(Tamilnadu).

2) The proposed scheme has within the substantial qualities of 1DCNN to achieve consistent and automatic extraction features which contributes towards the optimal prediction of soil nutrient composition.

3) This proposed enhanced 1DCNN model tries to address the limits which prevail in the available methods used for the process of prediction in terms of performance towards prediction, multifeatured processing capability, generalization, and prediction.

4) The proposed enhanced 1DCNN introduced layer level inner optimization and fine tuning towards the attainment of accurate soil nutrient(N,P,K) prediction.

5) Experiments of the proposed Enhanced-1DCNN DLM based soil nutrient prediction and fertilizer recommendation scheme is performed on the basis of metrics pertaining to increase in the rate of performance metrics as well as decrease in the error metrics by thus evaluating the advantages it poses in par with the baseline schemes used for the purpose of investigation.

6) The statistical and stability analysis was performed over the proposed model, which confirmed the stability of the utilized deep learning model for nutrient prediction and its respective fertilizer recommendation.

The other sections of this article are organized as drafted below. In the second section A comprehensive review was made on the available data mining and ML modelled soil-nutrient prediction schemes which have aided to the study of literature over the recent few years. The third section describes an inclusive view over the proposed Enhanced -1D CNN-scheme which predicts the soil nutrient composition of the soil dataset and recommends the suitable fertilizer adapting the layer-based feature optimization approach with appropriate validations. In the fourth section the proposed model's investigational results and its corresponding discussion pertaining towards predicting soil nutrient composition based on its metrics of performance and error measures in comparison with the baseline models has been discussed.

III. EXPERIMENT METHODOLOGY

The proposed Enhanced-one dimensional CNN based DLM for predicting the soil nutrient composition serves towards achieving sustainable groundnut cultivation through optimal prediction of soil nutrient composition and recommendation of required fertilizer need of the crop under cultivation based on the (N, P, K) input given. This Enhanced-1DCNN DLM adopted the significant parameters of various macro and micro nutrients along with soil pH in the course of predicting optimal soil nutrient composition as depicted in Fig. 1. Most commonly CNN based deep learning models are used to analyse the images, Deep two Dimensional Convolution neural networks, which might have several hidden layers and heaps of parameters, can learn objects of complex dimensions and patterns on Being trained on a large visual dataset with labelled values. This unique ability, when properly trained, defines to be the prime tool for several applications involving Two Dimensional signals such as imageries and frames of videos. However, this might not be a feasible choice in many applications involving one Dimensional signals, particularly if the training data is infrequent or confined towards any particular application. To gap this problem, 1D CNNs are proposed and it has quickly achieved the desired optimal performance levels in a variety of applications. Another striking benefit of 1D CNNs is that its configuration is simple and compact, which only perform 1D convolutions, which drives the route towards the usage of on demand cost feasible implementation of the hardware. Considering the impact of 1D CNNs in analysing temporal data, we have chosen to perform the process of soil nutrient prediction based on the optimized 1D CNN, which performs flattened layer level enhancement of conventional 2D CNNs.

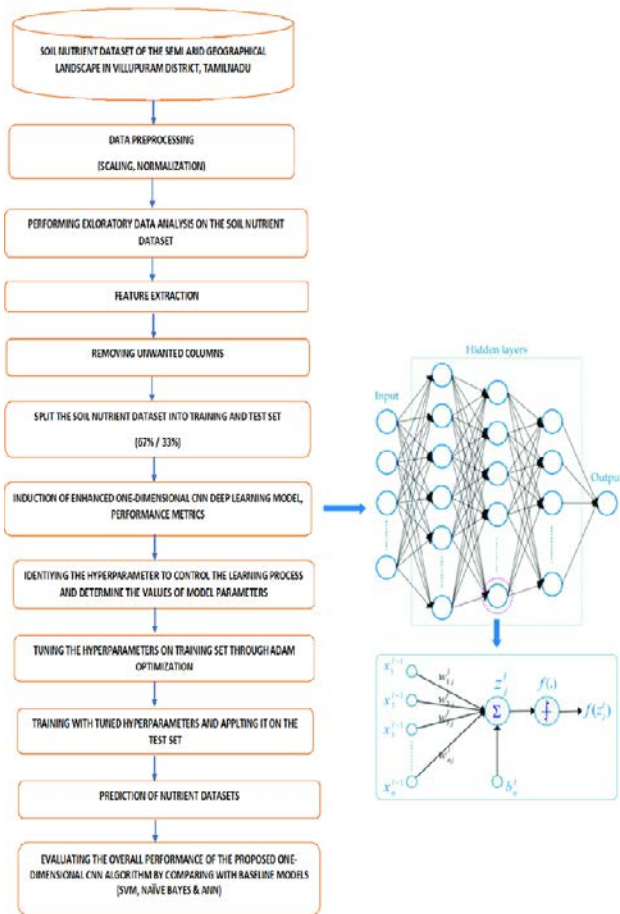


Fig. 1. Overall Process-flow Diagram for the Proposed Soil Nutrients Prediction Scheme using Enhanced 1D-CNN DLM.

A. Objective and Methodology

The cultivation of groundnut crop has been largely impacted by the improper or lack of proper assessment techniques used to exactly find the nutrient composition of the land under cultivation and needed percentage of correct fertilizers to be used based on the prediction scheme of the nutrients. The growth and the production yield of groundnut crops is largely impacted by the nutritious content of the soil, there are various set of macro and micro nutrients which contribute towards the proper growth of the groundnut crop starting from the sowing of seeds till towards the groundnut pods become mature enough to get harvested. The major set of nutrients present in the soil which contributes towards the crops growth are “Nitrogen[N], phosphorous[P] and Potassium [K]”, apart from other micro nutrient such as “Calcium (Ca), Sulphur[S], Zinc[Zn], Iron[Fe], Manganese [Mn] and Boron[B]”, soil parameters such as “pH, Soil electrical conductivity (EC), and Organic carbon (OC)” are also responsible in fixing out the overall fertility index of the soil. In this article we have chosen a semi-arid geographical landscape which is the most conducive soil pattern for the cultivation of groundnut and performed the model evaluation based on the nutrient datasets obtained from one such. The above given Table I, defines the parameters involved in the evaluation of the proposed Enhanced -1D CNN DLM.

TABLE I. PARAMETERS INVOLVED IN THE EVALUATION OF PROPOSED ENHC-1D CNN DLM

Geographical location of the dataset	Gingee (Taluk), Villupuram (District.), Tamilnadu (State), India
Latitude	12.2529° N
Longitude	79.4160° E
Source of Dataset	https://soilhealth.dac.gov.in/NewHomePage/NutriReport
Macro Nutrients	“Nitrogen[N], Phosphorous[P] and Potassium [K]”
other Nutrients	“Calcium (Ca), Sulphur[S], Zinc [Zn], Iron [Fe], Manganese [Mn] and Boron[B]”
Soil core parameters	“pH, Soil electrical conductivity (EC), and Organic carbon (OC)”
Nutrient classification classes on the datasets	Low, Medium & High

Excessive usage of artificial fertilizers, pollution variants and changing weather conditions make a huge impact by downgrading the quality of the soil under cultivation, and hence Prediction of the soil nutrient composition stays to be a vital factor towards framing up a well-defined precision farming prototype to uphold the soil fertility index and the crop production yield.

In this research work, a proposed 1D-CNN based Deep learning model has been chosen to train the nutrient datasets conducive for the sustainable cultivation of groundnut crop. This proposed 1D CNN deep learning model serves to the need of bridging the gap between the exact soil nutrient prediction for the cultivation of groundnut crop and its production rates through the improved optimal soil fertilizer recommendation schemes.

The one-dimensional CNN model can aggregate local features and lessen the data dimensions through convolutional learning and pooling operations. As a result, by repeatedly using convolution operations and pooling operations, a deep convolutional neural network can extract high-level features while significantly reducing the dimension of the output. In this article, raw nutrient dataset is used directly as the input to deep neural networks, as a result, the timing of the deep convolutional neural network output's high-level features is not disrupted.

B. Evaluation of the Proposed Model with Enhanced One-Dimensional CNN

A 1D CNN is a special case of a conventional neural network. Unlike traditional neural networks, in which the hidden layer is fully connected, a 1D CNN employs a unique network structure that alternates between the convolution and the pooling layers.

As shown in the above Fig. 2, the proposed 1D CNN has an input layer, three layers of convolutional (C_1, C_2, C_3), fully connected two layers ($F1, F2$) & an output layer. This 1D CNN stays to be an alternate enhanced version of 2D CNN.

These enhancements have proven to be more effective in certain applications which deal with 1D temporal data.

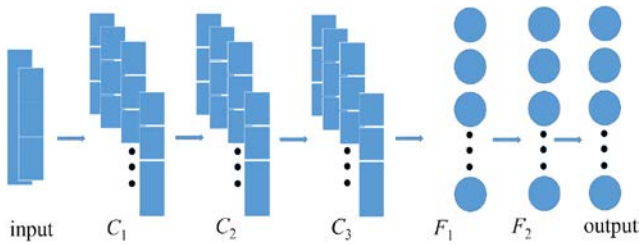


Fig. 2. Single CNN (1DS) Architecture.

Comparing with the 2D CNN, this 1DCNN are chosen to be more advantageous stating for the following reasons.

- The forward and the back propagation in one dimensional CNN require a simple array operation rather than a matrix operation.
- Due to its shallow Architecture, it involves in complex learning capability of 1D temporal data comparing with 2D CNN which has a deeper architecture, 1D-CNN are much convenient to train and use.
- The Complexity of involving more advanced hardware setup which includes the involvement of GPUs and cloud infrastructure in 2D-CNN is hugely reduced with respect to 1D-CNN where, general CPU implementation with a relatively fast processing speed makes 1D-CNN more opt for the usage.
- Compact 1DCNN have proven to be efficient in terms of performance pertaining to concise datasets.

As shown in the above Fig. 2, the 1D CNN Architecture consists of two distinct type of layers they are the “CNN-layers” in one Dimensional convolutions and pooling, and Fully-connected layers. The arrangement of a 1D-CNN is designed by the following hyper-parameters:

- The total levels of hidden CNN and fully connected layers (in our proposed Enhanced 1D-CNN model as shown in the above Fig. 2, there are 3 CNN layers and 2 fully connected layers).
- Defining the size for Filters in each CNN layer.
- Subsampling feature in each CNN layer.
- The choice of pooling and activation functions.

The input layer in 1D-CNN stays to be an inert layer which accepts the raw 1D temporal data as that of the conventional 2D-CNN. The output layer is a fully connected layer consisting of equal number of neurons as that of the number of classes. Fig. 3 depicts three consecutive CNN layers of 1D-CNN. As shown in the figure, 1D filter kernel has a size of 3 and a subsampling feature of 2. Here, the k th neuron in the CNN hidden layer l first performs a series of convolutions, the sum of which passes through the activation function f by a subsampling operation. In fact, this is the main difference between 1DCNN and 2DCNN, where the 1D array replaces the 2D matrix in both the kernel map and the feature map. Processing further, the raw 1D datasets are processed by the CNN layers and it starts learning to extract the features potential for the purpose of classification to be performed by

the Fully connected layers, as a result the process of feature extraction and classification are coupled together as a single process that which can be optimized so as to maximize the performance of the classification. This process of 1D CNN proves to be a major advantage as it results in the involvement of low computational complexity, since the only major operational cost is with the sequence of one dimensional convolutional which performs simply the linear weighted summing of two 1D arrays, which can be operated effectively during the forward and Back-Propagation operations in parallel.

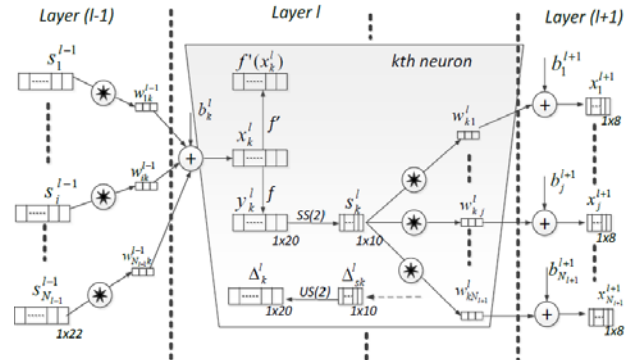


Fig. 3. Three Consecutive Hidden CNN Layers of a 1D CNN [33].

The procedure [33] of 1D Forward propagation (1D-FP) in each CNN-layer is shown in the equation (1) and the process of Back propagation is summed up as shown in the equation (2), is defined as follows:

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} \text{conv1D}(w_{ik}^{l-1}, s_i^{l-1}) \quad (1)$$

After the computation of weight and bias, they can be utilized to update the biases and weights with the learning factor, ϵ as,

$$w_{ik}^{l-1}(t+1) = w_{ik}^{l-1}(t) - \epsilon \frac{\delta E}{\delta w_{ik}^{l-1}}$$

and

$$b_k^l(t+1) = b_k^l(t) - \epsilon \frac{\delta E}{\delta b_k^l} \quad (2)$$

The process of forward and back-propagation in hidden 1D CNN layers is depicted below in the Fig. 4.

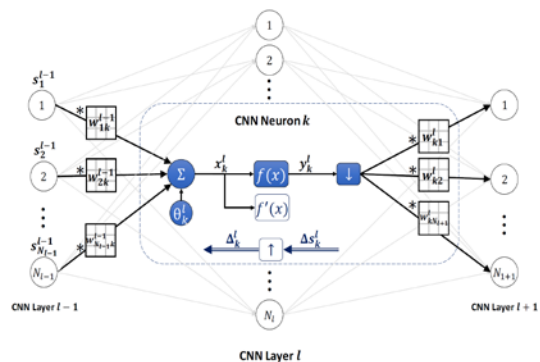


Fig. 4. CNN (Forward and Back-propagation in Hidden Layers).[33].

The process flow of the Back Propagation for the one-Dimensional temporal datasets in the training set can be stated as follows:

- 1) To initialize the weights and biases of the CNN.
- 2) For each Back Propagation iteration DO.
- 3) For each Nutrient Composition Value (Low, Medium, High) in the dataset, DO:
 - 4) FP: Forward propagation from the starting input level layer through the output layer to find the outputs of each neuron.
 - 5) BP: By Computing the delta error at the output level layer and to back-propagate it to first hidden layer to compute the delta errors,
 - 6) PP: To Post-process by computing the weight and bias sensitivities.
 - 7) Update: Updating the weights and biases by the (accretion of) sensitivities scaled up with the learning factor.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The evaluation of the proposed model using Enhanced 1D-CNN scheme and the benchmarked SVM, Naïve bayes and ANN schemes have been conducted based on the nutrient datasets of the Villupuram district obtained for a particular geographical region of Gingee. The dataset used for the current investigation is available online and it pertains to the years starting from 2016 to 2020. "Precision, recall score, F1-Measure, and Accuracy", as well as the error measures "RMSE, MAD, MAE, and R²" is calculated, and a comparison to the baseline models is derived through graphical comparisons for the proposed Enhanced ID-CNN. The datasets are mounted on Google Drive, and the model is evaluated on the Google Colab platform.

The training process of the proposed Enhanced ID-CNN structure is attained on the basis of Villupuram district nutrient dataset which are obtained from the online GIS servers, which are referred to the India soil Health System (soilhealth.dac.gov.in) maintained by the Indian Council of Agricultural Research. This web portal plays a prime role for providing soil related parameters across all the states in the country district wise. The soil nutrient constituents taken for the training purpose is obtained from the Tamilnadu state wise data of Villupuram district which spans over four years from (2016-2020). Data pre-processing is carried out before introducing into the model. The dataset is then partitioned based on the ratio of 67:33 for Training and Testing, respectively. Then the pre-processed data is inducted with the model evaluation of Enhanced 1D-CNN, during the process the model is evaluated for its performance measure and error metrics. The process of identifying the hyperparameters to control the learning process and to determine the values of model parameters is done by tuning the hyperparameters on the test set.

The proposed Enhanced 1D-CNN Deep learning model for predicting Soil nutrient composition for the sustainable

cultivation of the groundnut crop is a nascent Deep learning approach of this kind, hence for the purpose of evaluating the performance of the proposed Enhanced 1D-CNN model and its optimality consideration based on overall performance metrics and error measures it has been compared with standard and effective machine learning classifiers such as SVM, Naïve Bayes and ANN models and proved to be outperforming based on the evaluation results. The recommendation module evaluates the process of taking the given range of (N, P & K) values and generates the optimal fertilizers to be used for the sustainable cultivation of the groundnut crop. Since the general soil composition of the Villupuram district soil is Nitrogen Low by nature, the low predictions of Nitrogen below the minimum 17% of the basal need will have a nitrogen-based fertilizer recommendation for the sustainable cultivation of the groundnut crop, with respect to Phosphorous (P) based fertilizer recommendation, the higher percentage of phosphorous deposition found in Villupuram district soils along with the amount of needed phosphorous per hectare of groundnut cultivation demands a minimum of 35% of phosphorous requirement, which normally get satisfied due to its natural existence in the soil and hence rarely needs phosphorous based fertilizer recommendations. The requirement of Phosphorous is very much essential during the basal and flowering stages of the groundnut crops. The most important nutrient with respect to groundnut cultivation is the presence of Potassium(K), which is the most desirable nutrient required for the sustainable groundnut cultivation, starting from its early stage of growth till to its maturity, because it is responsible for making the crop disease resistive, regulates the water conditions within the plant cell, aids the crop in formation of proteins and chlorophyll and even often counterattacks the negative impacts of excess nitrogen supplements, supplied through fertilizers. To a minimum the percentage of Potassium will be nearly 55% in the total nutrient requirement for the sustainable cultivation of the groundnut crop, since the deposition of Potassium in the district of Villupuram seems to be in the Medium scale of level, fertilizers pertaining to Potassium are normally recommended when the predicted or the given values of Potassium is below the minimum scale required. The below Fig. 5, depicts the evaluation of the fertilizer recommendation module based on the given value of prediction.



Fig. 5. Fertilizer Recommendation Screen Shot.

A. Consideration of Performance Measures

- Accuracy: Adaptation of 1D-CNN has shown significant performance improvement in terms of the overall evaluation accuracy of the model in terms of classification. Comparative depiction of the accuracy of the proposed model is shown in Fig. 6.

$$ACCURACY = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

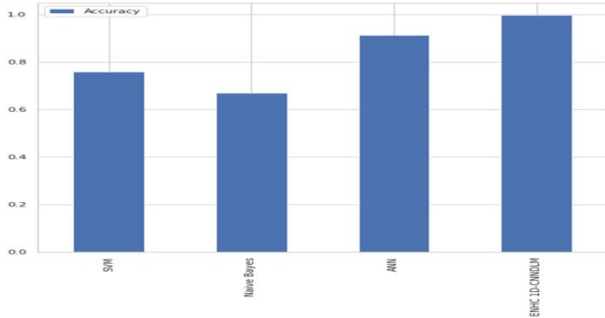


Fig. 6. Comparison of the Overall Prediction Accuracy of the Proposed Enhanced 1D-CNN.

- Precision: The overall Precision of the proposed Enhanced 1D-CNN described below is the ratio between the number of correctly identified positive predictions and the total number of positive predictions (True positive + False positive). Comparative depiction of the precision of the proposed model is shown in Fig. 7.

$$PRECISION = \frac{TP}{TP+FP} \quad (4)$$

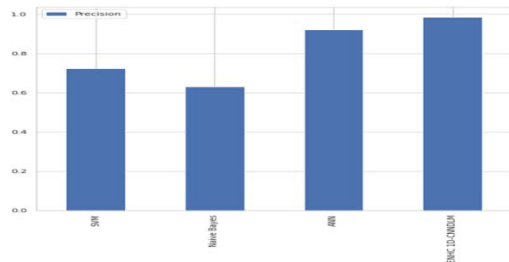


Fig. 7. Comparison of Precision Value for the Proposed Enhanced 1D-CNN.

Recall: As shown in Fig. 8, the recall value for the proposed Enhanced 1D-CNN is defined as the ratio of actually predicted true positive values to the overall sum of true predicted positive and false negative values.

$$RECALL = \frac{TP}{TP+FN} \quad (5)$$

F1-Measure: The harmonic mean of precision and recall is the F1 measure for the proposed Enhanced 1D-CNN, and its performance against the baseline model is depicted in Fig. 9 below.

$$F1 = 2 * \frac{PRECISION * RECALL}{PRECISION + RECALL} \quad (6)$$

e. Loss: This proposed Enhanced 1D-CNN Deep learning Model outputs a very minimal loss value by thus producing a more improved accuracy rate, The below Fig. 10 depicts it.

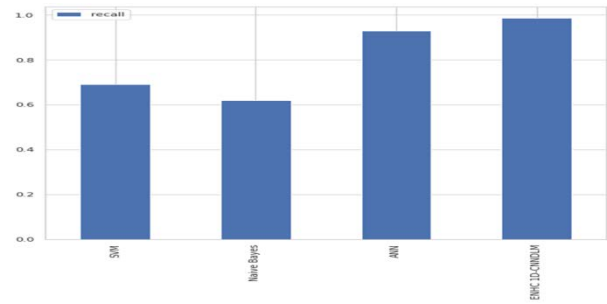


Fig. 8. Comparison over the Recall Score of the Proposed Enhanced 1D-CNN with Baseline Models.

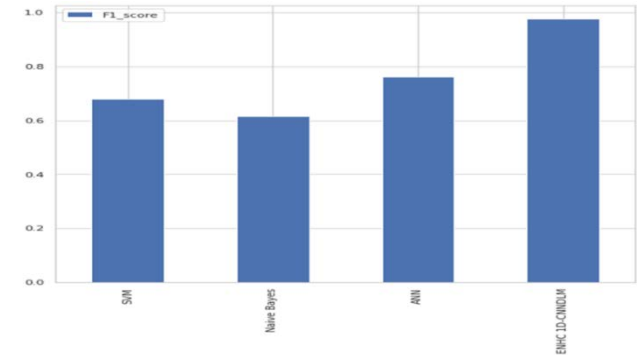


Fig. 9. Comparison over the F1-Measure of the Proposed Enhanced 1D-CNN with Baseline Models.

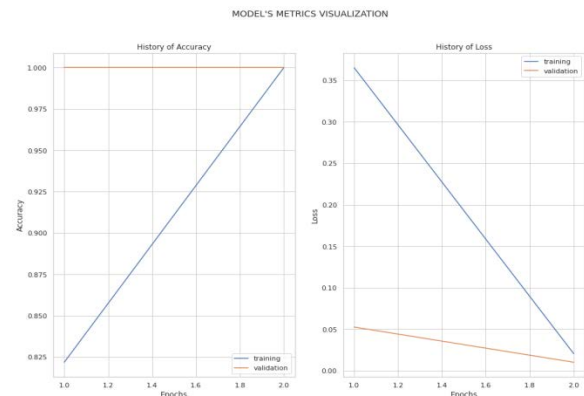


Fig. 10. Derived Accuracy and Loss of the Proposed Enhanced 1D-CNN.

The overall performance metrics of the proposed Enhanced 1D-CNN is depicted in Table II and its corresponding graphical representation is shown in the below Fig. 11.

TABLE II. PERFORMANCE METRICS OF THE PROPOSED ENHANCED 1D-CNN WITH THE BASELINE MODELS

	SVM	NB	ANN	ENHC ID-CNN/DLM
LOSS	0.23450	0.33670	0.9890	0.02347
ACCURACY	0.75894	0.66894	0.9124	0.99789
PRECISION	0.72345	0.62971	0.9197	0.98524
RECALL	0.69080	0.61987	0.9298	0.98741
F1_SCORE	0.67890	0.61620	0.7620	0.97745

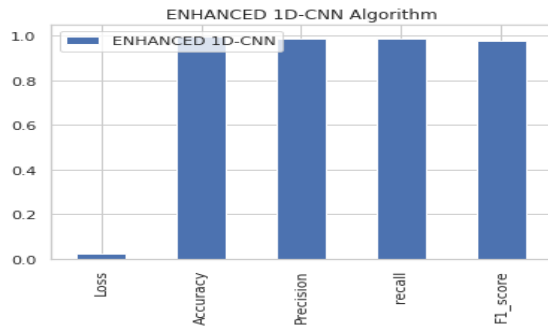


Fig. 11. Overall Performance Metrics of the Proposed Enhanced 1D-CNN.

Comparison of the Proposed Enhanced 1D-CNN with Baseline Models based on Performance measures, the below Fig. 12 depicts the optimality of the proposed model over other compared models.

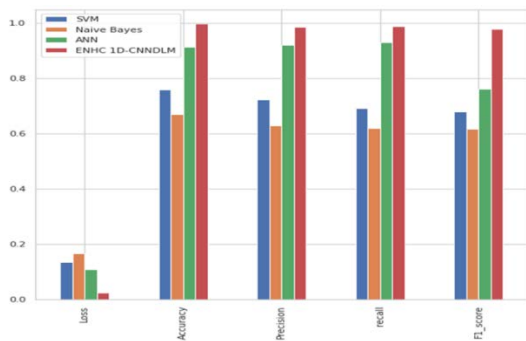


Fig. 12. Comparison of the Proposed Enhanced 1D-CNN with Baseline Models based on Performance Metrics.

B. Consideration of Error Metrics

The square root of the difference between the predicted values of the used model and the actual values associated with the study variable (soil nutrient) determined over the total number of observations is known as the Root of the Mean of the Squared Errors (RMSE) [32].

$$RMSE = \sqrt{\frac{ACT_{Val} - PR_{Val}}{Obs_{No}}} \quad (7)$$

MAE: The difference between the predicted values of the used model and the actual values associated with the study variable (soil nutrient) determined over the total number of observations is depicted by the mean of absolute values [32].

$$MAE = \frac{ACT_{Val} - PR_{Val}}{Obs_{No}} \quad (8)$$

Where, ACT_{Val} and PR_{Val} represents the actual and predicted values and Obs_{No} is the number of observations.

MAD: The average distance between each data point and the mean is the mean absolute deviation of a dataset. It gives us an idea of how variable a dataset is. [32].

$$MAD = \frac{\sum Absolute\ values\ of\ deviation\ from\ central\ measures}{Total\ no.\ of\ Observations} \quad (9)$$

The shown Table III depicts the calculated error metrics in par with all the baseline models with the Enhanced 1D-CNN and the corresponding graphical representation as shown in

Fig. 13 shows the performance of the Enhanced 1D-CNN with a minimal loss in the process of prediction and hence optimality in terms of achieving low error metrics. The calculated Coefficient of Determination (R^2) Value of the proposed Enhanced 1D-CNN scheme stands ahead in terms of evaluation with all the baseline models, and thus denotes how well the coefficient fits with the values in the training dataset.

TABLE III. COMPARISON OF THE ERROR METRICS COEFFICIENT OF DETERMINATION OF THE PROPOSED ENHANCED 1D-CNN WITH THE BASELINE MODELS

	SVM	NAÏVE BAYES	ANN	1D-CNN DLM
RMSE	0.8092	0.8872	0.7950	0.4839
MAE	0.6549	0.7872	0.6321	0.2342
MAD	0.3238	0.6871	0.2871	0.2270
R²	0.6123	0.5321	0.9312	0.9872

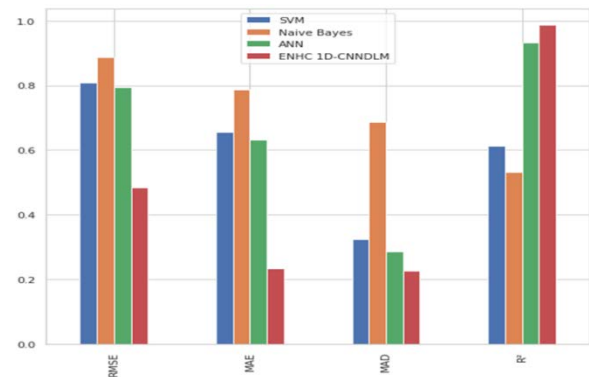


Fig. 13. Error Metrics and Coefficient of Determination (R^2) Comparison of the Proposed Enhanced 1D-CNN with Baseline Models.

The proposed Enhanced 1D-CNN schemes' prediction efficiency is tested using the Villupuram district soil nutrient dataset by deriving the confusion matrix (see Table IV (A, B, C)). The computed confusion matrix for the proposed model, has shown evident prediction results by producing most optimal prediction accuracy rates for predicting the major macro nutrients of the soil Nitrogen (N), Phosphorous (P) and Potassium (K), based on the classification sets Low, Medium and High. This classification labels are defined on the basis of the nutrient dataset values, which represents the exact soil nutrient composition of these nutrients in the district of Villupuram.

The proposed Enhanced 1D-CNN model and the compared benchmarked schemes go through a three-step experimental process. Initially, the proposed Enhanced 1D-CNN model along with its benchmarked comparative models are compared for the performance measures such as (Accuracy, Precision, Recall, and F1-score) as well as error measures (RMSE, MAE, MAD, MAE). The model is also evaluated for its coefficient of determination value (R^2) pertaining to the Villupuram district soil nutrient dataset. Figure 11 shows the performance of the proposed Enhanced 1D-CNN scheme and the benchmarked SVM, Nave Bayes, and ANN models in terms of mean accuracy and precision. The proposed Enhanced 1D-CNN deep learning architecture ensured maximum overall accuracy and precision, as the speed with which it deduced

robustness features that could potentially influence the prediction of soil Nutrient Composition was way ahead when compared with other learning models. As a result, the proposed Enhanced 1D-CNN scheme has the potential to improve accuracy by 23.89 per cent, 32.89 per cent, and 8.54 per cent over the baseline models SVM, Nave Bayes, and ANN, respectively.

TABLE IV. DERIVATION OF CONFUSION MATRIX DEPICTING THE PREDICTING EFFICIENCY OF (N, P, K) VALUES BY THE PROPOSED ENHANCED 1D-CNN, (A) CONFUSION MATRIX FOR PHOSPHOROUS CLASSIFICATION, (B) CONFUSION MATRIX FOR NITROGEN CLASSIFICATION, (C) CONFUSION MATRIX FOR POTASSIUM CLASSIFICATION

(A)

		PREDICTED PHOSPOROUS (P) COMPOSITION			
		LOW	MEDIUM	HIGH	Σ(LMH)
ACTUAL PHOSPOROUS (P)	LOW	378	1	0	379
	MEDIUM	0	592	2	594
	HIGH	7	0	580	587
	Σ(LMH)	385	593	582	1560

(B)

		PREDICTED NITROGEN (N) COMPOSITION			
		LOW	MEDIUM	HIGH	Σ(LMH)
ACTUAL NITROGEN (N) COMPOSITION	LOW	1260	1	0	1261
	MEDIUM	0	260	1	261
	HIGH	7	0	31	38
	Σ(LMH)	1267	261	32	1560

(C)

		PREDICTED POTASSIUM (K) COMPOSITION			
		LOW	MEDIUM	HIGH	Σ(LMH)
ACTUAL POTASSIUM (K) COMPOSITION	LOW	624	1	0	625
	MEDIUM	0	770	2	772
	HIGH	7	0	156	163
	Σ(LMH)	631	771	158	1560



Fig. 14. Heat Map Depicting the Correlations of the Soil Dataset.

The correlation between the variables is shown in Fig. 14 using a heat map. Correlation matrix that shows the relationship between two parameters in a dataset. The heat map is used in exploratory data analysis to check for correlations between the data. The heat map is used to visualize the parameter correlation matrices as well as to determine which parameters influence the output variable. The level of correlation between the various macro and micro nutrients, as well as pH, Electrical Conductivity (EC), and organic carbon (OC), has been absorbed in the heat map shown above. There is a strong correlation between the macro nutrient (Nitrogen) and the micro nutrient (Calcium), which is essential for the growth of groundnut crops during the time of pod formation, as shown in Fig. 14, but there are also instances of weak correlation. The value of heat map ranges from +1 to -1 where, Positive values indicate positive correlation, while negative values indicate negative correlation. A stronger linear association exists between data points that are closer together, whereas a weaker linear association exists between values that are closer to zero.

V. CONCLUSION

The prime reason for the loss of soil quality is due to the improper soil and crop management strategies deployed in the process of farming. Excessive usage of chemical fertilizers without the exact knowledge of required nutrients for the cultivation of crops has led to the gradual decline in the soil quality which has contributed towards the gradual decline in the production yield of the crops. Groundnut Crops being the most predominant oilseed crop cultivated in the state of Tamil Nadu, its cultivation has seen too such impacts due to the excessive fertilizer usage and improper cultivation patterns without understanding the actual nutrient of the soil, one of the major causes for the fluctuation seen in groundnut pod growth patterns and production, is the differing soil nutrient compositions of the land which is under cultivation. In this research article we proposed a novel deep learning based approach adapting Enhanced 1D-CNN scheme, towards predicting the soil nutrient composition for the chosen soil nutrient dataset pertaining to the geographical landscape of Villupuram district in the state of Tamil Nadu. The experimental results carried out has clearly shown the effectiveness of the proposed Enhanced 1D-CNN DLM, in terms of the overall performance measures (Accuracy, Precision, Recall, F1 Score, Recall and a very minimal Loss value), the optimality of the proposed model has been compared for its effectiveness with baseline models such as (SVM, Naïve Bayes and ANN) and proven to be outperforming all the baseline models in terms of increased performance measures resulting in the overall prediction accuracy of 99.78% and very minimal error measures. The fertilizer recommendation based on the predicted nutrient composition of the soil makes this proposed model more productive in terms of addressing the objective of this research work. This proposed scheme for the sustainable cultivation of groundnut crop may be considered as a reference scheme for crop specific precision farming moving further.

REFERENCES

- [1] S.NIKAM. (2015). A Comparative study of classification Techniques in Data Mining Algorithms. *oriental journal of computer science & technology*, 8(1), 13-19.
- [2] Jay Gholap, Anurag Ingole., Shailesh Gargade, Jayesh Gohil, & Vahida Attar. (2012). Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction. *ResearchGate-2012*.
- [3] Adhikary, T., Das, A. K., Razaque, M. A., Chowdhury, M. E., & Parvin, S. (2015). Test implementation of a sensor device for measuring soil macronutrients. 2015 International Conference on Networking Systems and Security (NSysS). doi:10.1109/nsyss.2015.7042951.
- [4] Ahmad, S., Kalra, A., & Stephen, H. (2010). Estimating soil moisture using remote sensing data: A machine learning approach. *Advances in Water Resources*, 33(1), 69-80. doi:10.1016/j.advwatres.2009.10.008.
- [5] Chiranjeevi M N, Ranjana B Nadagoudar(2018)Analysis of Soil Nutrients using Data Mining Techniques. *International Journal of Recent Trends in Engineering and Research*, 4(7), 103-107. doi:10.23883/ijrter.2018.4363.pdt1c.
- [6] Antonios Morellos a, Xanthoula-Eirini Pantazi a, Dimitrios Moshou a, Thomas Alexandridis c, Rebecca Whetton b, Georgios Tziotzios a, & Jens Wiebensohn d. (2016). Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectros- copy. Published by Elsevier. doi:10.7287/peerj.5714v0.2/reviews/1.
- [7] Balducci, F., Impedovo, D., & Pirlo, G. (2018). Machine Learning Applications on Agricultural Datasets for Smart Farm Enhancement. *Machines*, 6(3), 38. doi:10.3390/machines6030038.
- [8] Bhargavi, P., & Jyothi, S. (2009). Applying Naive Bayes Classification Technique for Classification of Improved Agricultural Land soils. *IJCSNS International Journal of Computer Science and Network Security*, VOL.9 No.8, August 2009, 9(8), 117-122.
- [9] Dr. K. Arunesh, & V. Rajeswari. (2017). agricultural soil lime status analysis using data mining classification techniques. *International Journal of advanced technology in engineering and science*, 5(2), 27-35. Retrieved from www.ijates.com.
- [10] Ingale, V., Vaidya, R., Phad, A., & Shingare, P. (2016). A sensor device for measuring soil macronutrient proportion using FPGA. 2016 International Conference on Communication and Signal Processing (ICCSP). doi:10.1109/iccsp.2016.7754237.
- [11] Kulkarni, K., & Vijaya, P. A. (2017). A comparative study of land classification using remotely sensed data. 2017 International Conference on Computing Methodologies and Communication (ICCMC). doi:10.1109/iccmc.2017.8282720.
- [12] Leisa J. Armstrong, Dean Diepeveen, & Rowan Madder. (n.d.). The application of data mining techniques to characterize agricultural soil profiles.
- [13] Mathavi Parvathi, T. (2017). automated soil testing process using combined mining patterns. *best: International Journal of Humanities, Arts, Med- icine and Sciences (BEST: IJHAMS)*, 5(8), 76-79.
- [14] Mrs. N. Hemegeetha, & Dr. G.M. Nasira2. (2016). Analysis of Soil condition Based on pH value Using Classification Techniques.
- [15] Panchamurthi, S. (2019). Soil Analysis and Prediction of Suitable Crop for Agriculture using Machine Learning. *International Journal for Research in Applied Science and Engineering Technology*, 7(3), 2328-2335. doi:10.22214/ijraset.2019.3427.
- [16] Patil, T. R., Shamshuddin K, Patil, R., & Sadanand P. (2016). Krushi Samriddhi: A decision support system for farmers to get high crop yield. 2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT). doi:10.1109/ic- ctict.2016.7514548.
- [17] Pawar, M., & Chillarge, G. (2018). Soil Toxicity Prediction and Recommendation System Using Data Mining In Precision Agriculture. 2018 3rd International Conference for Convergence in Technology (I2CT). doi:10.1109/i2ct.2018.8529754.
- [18] Puno, J. C., Sybingco, E., Dadios, E., Valenzuela, I., & Cuello, J. (2017). Determination of soil nutrients and pH level using image processing and artificial neural network. 2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM). doi:10.1109/hnicem.2017.8269472.
- [19] Rahman, S. A., Chandra Mitra, K., & Mohidul Islam, S. (2018). Soil Classification Using Machine Learning Methods and Crop Suggestion Based on Soil Series. 2018 21st International Conference of Computer and Information Technology (ICCIT). doi:10.1109/iccitech.2018.8631943.
- [20] Rajeswari, V., & Arunesh, K. (2016). Analysing Soil Data using Data Mining Classification Techniques. *Indian Journal of Science and Technol- ogy*, 9(19). doi:10.17485/ijst/2016/v9i19/93873.
- [21] Ramesh, V., & Ramar, K. (2011). Classification of Agricultural Land Soils: A Data Mining Approach. *Agricultural Journal*, 6(3), 82-86. doi:10.3923/aj.2011.82.86.
- [22] Regalado, R. G., & Dela Cruz, J. C. (2016). Soil pH and nutrient (Nitrogen, Phosphorus and Potassium) analyzer using colorimetry. 2016 IEEE Region 10 Conference (TENCON). doi:10.1109/tencon.2016.7848458.
- [23] JuhiReshma S R, Dr. D. John Aravindhar (2021). Fertilizer Estimation using Deep Learning Approach. *Nat.Volatiles&Essent.Oils,2021;8 (4):5745-5752*.
- [24] Anguraj.K, Thiyaneswaran.B, Megashree.G, Preetha Shri.J.G, Navya.S, Jayanthi. J Crop(2021). Recommendation on Analysing Soil Using Machine Learning, *Turkish Journal of Computer and Mathematics Education* ,Vol.12 No.6 (2021), 1784-1791.
- [25] apeda.gov.in/apedawebsite/HACCP/2018_Groundnut_Survey_Report.pdf.
- [26] agritech.tnau.ac.in/agriculture/agri_seasonandvarieties_groundnut.html.
- [27] apeda.gov.in/apedawebsite/SubHead_Products/Ground_Nut.htm.
- [28] <https://www.ikisan.com/tn-groundnut-history.html>.
- [29] Soilhealth.dac.gov.in/PublicReports/MacroNutrientsStateWiseNPK.
- [30] Efficiency Of Groundnut Cultivation in Tamil Nadu, Dr. S. Sudhamathi, *Science Technology and Development Volume VIII Issue XII DECEMBER 2019 ISSN: 0950-0707, Page No: 361*.
- [31] R. Jayalakshmi, M. Savitha devi, Predictive Model Construction for Prediction of Soil Fertility using Decision Tree Machine Learning Algorithm, *INFOCOMP*, v. 20, no. 1, p. 49-55, June, 2021.
- [32] Sivasankaran, S., K. Jagan Mohan, and G. Mohammed Nazer. "Soil Moisture Quantity Prediction using Optimized Deep Learning Supported model for Sustainable cultivation of Groundnut plant." *NVEO-NATURAL VOLATILES & ESSENTIAL OILS Journal* NVEO (2021): 13184-13205.
- [33] S. Kiranyaz, A. Gastli, L. Ben-Brahim, N. Alemadi, M. Gabbouj, Real-Time Fault Detection and Identification for MMC using 1D Convolutional Neural Networks, *IEEE Trans. Ind. Electron.* (2018). doi:10.1109/TIE.2018.2833045.
- [34] Dietterich, Thomas G. "Approximate statistical tests for comparing supervised classification learning algorithms." *Neural computation* 10.7 (1998): 1895-1923.