Indonesian Journal of Electrical Engineering and Computer Science
Vol. 25, No. 3, March 2022, pp. 1688~1695
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v25.i3.pp1688-1695

An artificial intelligence solution for crop recommendation

Varshitha D. N., Savita Choudhary
Department of Computer Science and Engineering, Sir M Visvesvaraya Institute of Technology, Bengaluru, India

Article Info

Article history:

Received Sep 20, 2021
Revised Dec 9, 2021
Accepted Jan 20, 2022

Keywords:

Deep learning
Deep neural network
Machine learning classifiers
NPK
Prediction

ABSTRACT

Agriculture is the major occupation in India. The development of India is in the hands of farmers. Farmers are said to be our nation's backbone, so there is a need to support our farmers technologically so that the difficulties of traditional agricultural practices would be overcome and also there will be positive impact on the yield, harvest, healthy crop output and the income of the farmers. Farmer needs awareness about his soil and the methods to improve his soil to grow the healthy crops. We propose an approach which involves deep learning and some IoT features to help our farmers. Soil parameters such as nitrogen, phosphorous, potassium (NPK), pH, organic carbon, moisture content and few more things are considered for predicting the fertility of the soil and also to predict the right crops to be grown and nutrition required for it. We have developed a deep neural network model to predict the crop which can be suitably grown in the soil. We have also implemented the other machine learning classifiers on the same collected dataset to test the accuracies of each classifier and our deep neural network model.

Corresponding Author:

Varshitha D. N.
Department of Computer Science and Engineering, Sir M Visvesvaraya Institute of Technolgy
Bengaluru, Karnataka, India
Email: varsha.swara@gmail.com

1. INTRODUCTION

Majority of rural India is still dependent on agriculture for the livelihood; also agriculture is the biggest sector of economy. Though India has revolutionized in the area of agriculture, still there is scope for improving the methods of agriculture and crop yield enhancement using more scientific and innovative approaches. A lot of researches happen in the area of agriculture every day. Farmers still resist to apply modern techniques for agriculture because of difficulty faced in getting adopted to new approaches [1], [2]. If the approaches are easily accessible with cost and time efficient methods, then the number of farmers switching to modern techniques from tradition way will be more. Agriculture is not only the main sector of economy, it also provides food to people and gives the raw materials to industries. The growing demand to provide food also encourages to improve the agricultural methodologies

A lot of farmers still need awareness about the soil to promote healthy crop growth and to increase the yield and income. Soil is the vital component of agriculture. Information regarding the soil like fertility, estimated yield, lacking components in the soil, the crops which can be grown from the soil and many more soil related things would be beneficial to all the farmers to choose the correct crops which can be grown in their land, farmer should know which crop's growth is facilitated and which are not so that unpredictable circumstances are avoided after sowing the seeds. If farmer grows the crops by considering his economic conditions, soil parameters and available facilities, it would turn into a boon for him in future as he can expect healthy crop growth, more yield and better income. So, if farmers are guided in this right direction using modern technologies, well-being of the farmers are assured and country's growth is also possible.

Soil has several factors or parameters in it, the proper composition of each and every parameter would help in growth of plants. Major components are macronutrients like NPK and micro nutrients such as zinc, iron, manganese, and magnesium and other factors like pH level, moisture content, temperature, electric conductivity, organic carbon and external factors like climatic conditions and many more factors, all together facilitate the crop growth [3]-[6]. The technique like artificial Intelligence utilising the data of soil parameters can be used to predict fertility rate, recommend crops and nutrition in minimum time and cost. This would make farmer convenient to know about his soil and choose the better crop. Also, the suggested nutrients or fertilizers would help to enhance the crop growth and yield [7]-[11].

Deep neural network, the part of deep learning is a branch of artificial intelligence and sub branch of machine learning. Neural network works like a human brain and has capability to learn from the data. Deep neural network has many layers in the network. Each layer contains multiple neurons and neurons are connected to each other, Neurons process the information and pass it on to the next layer. Activation function and weights are responsible for the strengths of the signal. Deep neural networks can solve various modern day problems and produce accurate results [12]. Robustness, parallel computation, self-learning ability and flexibility are the major pros of neural network. If good amount of data is provided to deep neural network then classification and regressions are performed well hence it is a great solution for predicting fertility and suggesting the right crops and nutrients [8], [12]-[15]. We have discussed the solution for crop prediction which is a part of our 'soil fertility and crop friendliness detection and monitoring system' in the subsequent sections of our paper which includes architecture, methodology and results of the proposed work. The prediction of crops is a multi-class classification problem where a category or a particular class is identified as a result. There are several classifiers present, out of which we have chosen few well known classifiers to test and compare against the accuracy of our deep neural network model.

2. PROPOSED ARCHITECTURE

The 'soil fertility and crop friendliness detection and monitoring system' is implemented using deep neural networks and machine learning techniques to predict the fertility rate of the soil and also to recommend the right crops which can be grown in that soil. After suggesting the right crops, the right fertilizer or nutrients needed for the suggested crops are also recommended. The proposed architecture has soil classifier, comparison module, prediction module, crop recommender and nutrition recommender modules as illustrated in the Figure 1.

If the farmer knows about the crops which can be grown in his land, he can input the same to the system or else the soil sensors can be used to capture the soil parameters like NPK, pH level, moisture, temperature, electric conductivity, organic carbon level and so on. The input is next transferred to soil classifier module, along with user input to classify and move the control to desired module. If the Soil type is known, AI comparison module will work with the help of pre stored Nutrition dataset which is comprised of soil parameters, crops, fertility rate or sensor reading. This module predicts fertility rate of the soil for the user known crop. If the farmer does not have any idea about the type of soil, the classifier module would transfer the control to AI prediction module to predict the fertility rate and the best crop which can be grown in his soil with the help of Nutrition dataset. If farmers migrate from one place to another or if farmers want to grow new crop in their land then crop prediction will be really beneficial. The architecture of the prediction module is shown in Figure 2.

The crop recommender module along with the crop data set which has soil parameters and the crops grown data, would suggest the other four to five crops which could be grown using the given soil sample. The nutrition recommender module would take up the output of crop recommender system and along with Fertilizer dataset, suggests the required nutrients or fertilizers for enhancing fertility that will help to grow each crop suggested. The fertility rate of the soil, the crop recommendations and the nutrition or fertilizer suggestions would be stored in cloud and sent in the form of soil reports to the farmers to their phones in easily readable and understandable format.

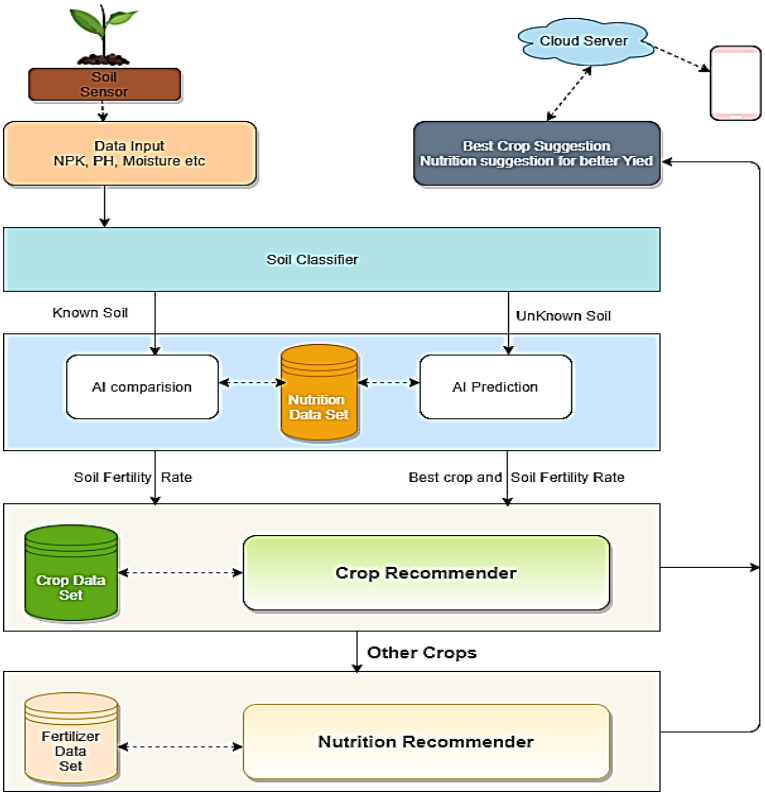


Figure 1. Architecture of the 'Soil fertility & crop friendliness detection & monitoring system'

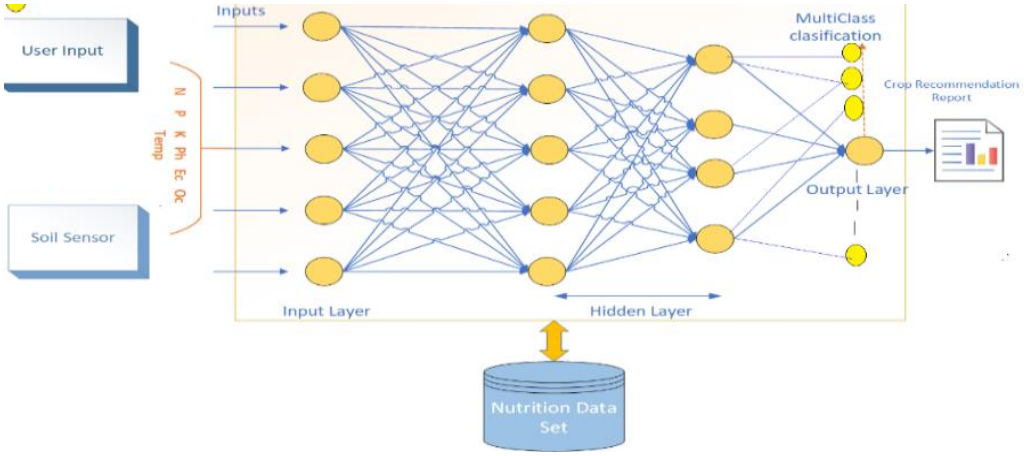
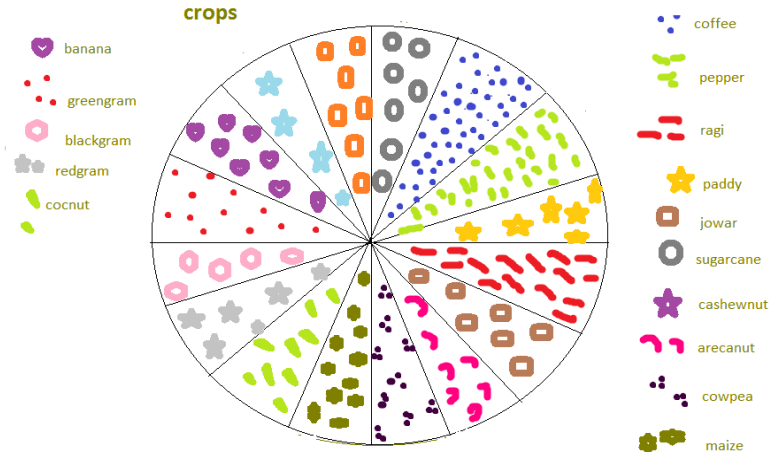


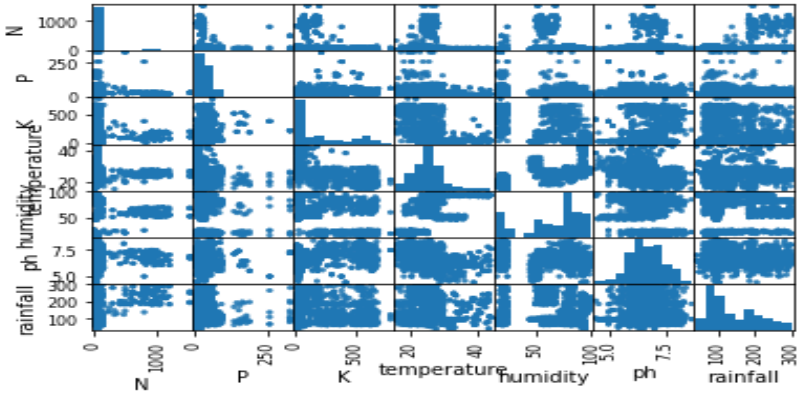
Figure 2. Architecture of AI prediction module

3. DATASET

With the help of the website of Indian Government and soil testing sensors, data has been collected from some regions of Mysore district and Kodagu district of Karnataka. The crops chosen are paddy, coffee, maize, cowpea, red gram, banana, groundnut, areca nut, coconut, jowar, green gram, ragi, black gram, pepper, cashew nut, sugarcane. There are 16 crops collected and chosen having 300 entries for each crop hence total of 4800 rows of data is present. Figure 3(a) represents the Nutrition dataset and its distribution. The soil parameters considered are nitrogen (N), phosphorous (P), potassium (K), temperature, humidity, pH, rainfall, electric conductivity and organic carbon. The crop column is considered as the target since the crop label or class has to be predicted from the soil parameter values. The dataset scatterplot matrix represents the relationship between soil parameters as shown in the Figure 3(b).



(a)



(b)

Figure 3. Dataset representation in (a) crops distribution in collected nutrition dataset and (b) dataset scatterplot matrix

4. METHODOLOGY & RESULTS

4.1. Deep neural network model

In this paper we have discussed more on AI prediction module to predict the best crop. We have used keras deep learning neural network for building this module [16]. Keras model is flexible and easily deployable in nature. Keras sequential model is a simple network model used for classifying the multiple crops where Multi class classification is done for categorizing the crop into the particular class and predicting the right crop.

Google Colab platform is used for implementing the model, important libraries like Scikit-learn, Numpy, and Pandas, necessary classes and functions are loaded and imported. Dataset which is collected comprising of 4800 rows has been stored in CSV format and loaded into our Colab file. Data shuffling and normalization is done as initial steps to avoid biasing and redundancy and to increase the accuracy of the model. Dataset is split into 80% training data and 20% testing and validation dataset. Since the output is categorical, it is encoded, later the neural network and its layers are defined. Considering the number of soil parameters which are meant for input, input dimension is defined. With Relu activation function the input layer and hidden layers are built. Softmax activation function is used at output layer, we have used Adam optimizer for higher efficiency and loss function with categorical cross entropy is used. Keras classifier is passed to fit function to train the model and run with 100, 150 200 and 500 epochs, the loss and accuracy is obtained and analysed and later prediction is performed. The summary of result after 500 epochs is represented in Figure 4(a) and crop prediction instance having predicted values against the real values are shown in Figure 4(b). Looking at the loss curve graph as shown in the Figure 5, we can say the learning rate is high with negligible difference between training and validation loss points and the model is a good fit. The Keras deep learning model implemented and run for the collected dataset gives the accuracy of classification approximately 87%.

summary of the result after each epoch:

	loss	accuracy	val_loss	val_accuracy	epoch
495	0.197731	0.867708	0.200129	0.854167	495
496	0.198495	0.869531	0.223541	0.862500	496
497	0.193396	0.874479	0.218459	0.835417	497
498	0.195570	0.868490	0.222443	0.852083	498
499	0.197824	0.869531	0.213771	0.868750	499

(a)

	crop	N	P	K	temperature	humidity	ph	rainfall
paddy	paddy	90	42.00	43.00	20.8797	82.0027	6.50299	202.936
maize	maize	61	21.00	123.00	25.1421	65.2619	6.2	76.6846
pepper	coffee	91	21.00	26.00	26.3338	57.3647	7.26131	191.655
ragi	ragi	0	18.20	127.00	18.153	19.386	6.8	194.577
blackgr...	blackgram	0	27.00	558.00	23.38	21.9888	6.9	182.904
arecanut	arcenut	70	5.36	5.50	39.7315	91.1222	6.73	122.763

(b)

Figure 4. Results of our neural network model: (a) result summary of our model and (b) an instance of crop prediction

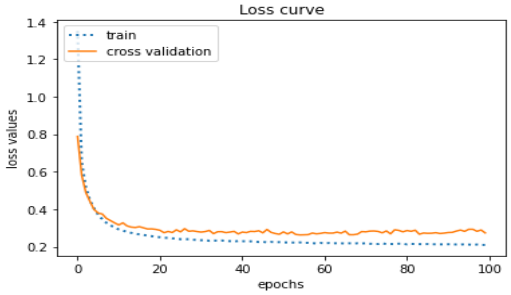


Figure 5. Loss curve graph of our model

4.2. The other machine learning classifiers implemented are

4.2.1. K-nearest neighbor (KNN)

KNN is the simplest algorithm used for classification of multiclass dataset which can be also used for crop recommendation [17]-[20]. The similarity between the available data and new data is assumed and the new data is put into the category which is most similar to the available category. The 'K' has to be selected first, then it requires computation of distance between data points or K neighbors using method like Euclidian distance. The KNN classifier yields the accuracy of 64.7% for our dataset.

4.2.2. Decision tree classifier

Decision tree classifier is a well know ML classifier which can be also used for predicting crops [21], [22]. A Tree is constructed using recursive binary splitting method. The tree is iteratively split until lowest subsets are obtained. Whenever is new data is considered for categorizing, trail of tests which is arranged in hierarchical way are performed to obtain the class label. The decision tree classifier yields the accuracy of 70% for our dataset.

4.2.3. Support vector machine (SVM)

SVM is best suited multi class classification for predicting crops [17], [20], [23]. A Hyperplane is created from SVM algorithm by choosing the vector points which will be the best decision boundary, this hyperplane is used for taking decision and segregate the data points to proper category where it belongs to. SVM is an efficient algorithm which works well in high dimensional space. The Support vector classifier yields the accuracy of 71.56% for our dataset.

4.2.4. Gaussian NB

Naïve Bayes (NB) is a well know ML classification method used for crop recommendation [17], [18], [22], [23]. Gaussian NB is special kind of Naïve Bayes classification method where features are

assumed to have normal distribution. Naïve Bayes works using Bayes theorem where conditional probability is calculated. Gaussian NB is a faster algorithm to categorize high dimensionality data. The Gaussian NB classifier yields the accuracy of 72% for our dataset.

4.2.5. Linear discriminant analysis (LDA)

Linear discriminant analysis can be used for well-known classification problems in agricultural domain [24], [25], LDA is used for classification as well as dimensionality reduction. LDA is a simple classification algorithm which finds linear combination of characteristics to categorize into two or more classes hence it is a preferred linear classification technique. LDA divides the classes into two or more groups by showing spaces in higher dimensions and in lower dimension. The Linear Discriminant Analysis classifier yields accuracy of 60.9% for our dataset. The results having confusion matrix of the implemented classifiers are represented in Figures 6 and 7. The comparison of accuracies of deep neural network and various classifiers are shown in the Figure 8.

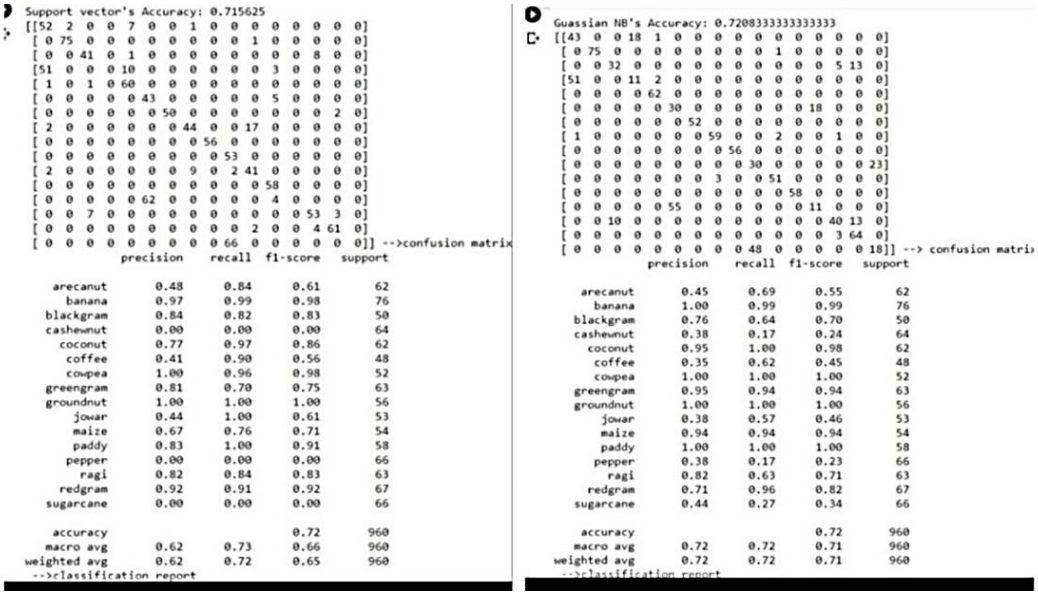


Figure 6. Classification report of SVM & Gaussian NB classifiers

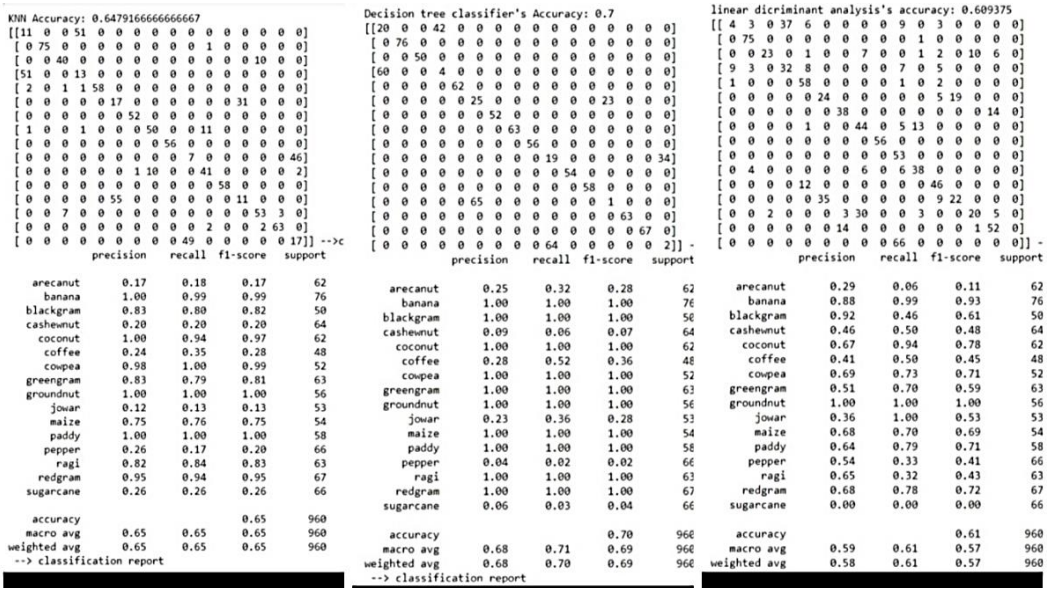


Figure 7. Classification report of KNN, decision tree and LDA classifiers

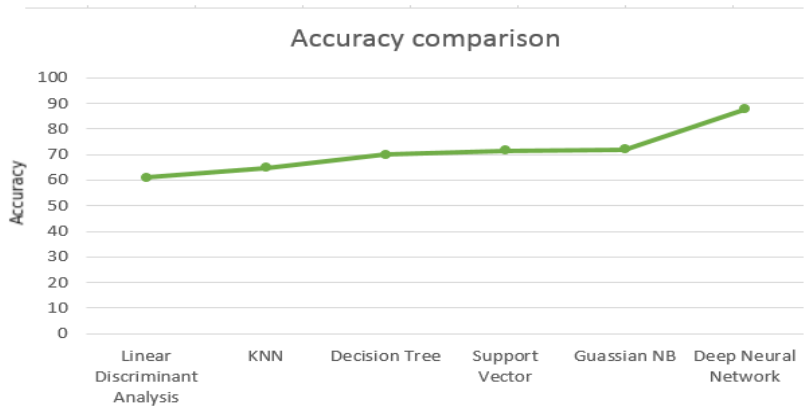


Figure 8. Accuracy comparison of all the implemented ML classifiers and deep learning model

5. CONCLUSION

We have implemented deep neural network and many machine learning algorithms as classifiers for our collected Nutrition dataset and obtained the accuracies of our neural network model and other classifiers. Our deep learning model has 87% accuracy and we obtained 71.5%, 72%, 65%, 70%, and 61% accuracy for SVM, Gaussian NB, KNN, decision tree and LDA respectively. So, the deep neural network has highest accuracy among all our algorithms and models with added advantages of its self-learning ability, robustness, flexibility and many more things. Classification is 87% accurate which is really a better result for predicting the right crops and help out the farmers to choose the most appropriate crop for his land based on the soil properties.

REFERENCES

[1] K. Katke, "Precision Agriculture Adoption: Challenges of Indian Agriculture," *International Journal of Research and Analytical Reviews*, vol. 6, no. 1, pp. 863-869, 2019.
[2] S. M. Dev, "Small farmers in India: Challenges and opportunities," *Indira Gandhi Institute of Development Research, Mumbai Working Papers 2012-014*, 2014.
[3] G. Warudkar and S. Dorle, "Review on sensing the fertility characteristics of agriculture soils," *2016 International Conference on Information Communication and Embedded Systems (ICICES)*, 2016, pp. 1-6, doi: 10.1109/ICICES.2016.7518850.
[4] M. Masrie, M. S. A. Rosman, R. Sam, and Z. Janin, "Detection of nitrogen, phosphorus, and potassium (NPK) nutrients of soil using optical transducer," *2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, 2017, pp. 1-4, doi: 10.1109/ICSIMA.2017.8312001.
[5] A. A. Chitragar, S. M. Vasi, S. Naduvinamani, A. J. Katigar, and T. I. Hulasogi, "Nutrients detection in the soil: review paper," *International Journal on Emerging Technologies (Special Issue on ICRIET-2016)*, vol. 7, no. 2, pp. 257-260, 2016.
[6] A. Amrutha, R. Lekha, and A. Sreedevi, "Automatic soil nutrient detection and fertilizer dispensary system," *2016 International Conference on Robotics: Current Trends and Future Challenges (RCTFC)*, 2016, pp. 1-5, doi: 10.1109/RCTFC.2016.7893418.
[7] D. Vadalia, M. Vaity, and K. T. D. Kapse, "Real Time soil fertility analyzer and crop prediction," *International Research Journal of Engineering and Technology (IRJET)*, vol. 4, no. 3, 2017.
[8] T. V. N. Rao and S. Manasa, "Artificial Neural Networks for Soil Quality and Crop Yield Prediction using Machine Learning," *International Journal on Future Revolution in Computer Science & Communication Engineering*, vol. 5, no. 1, 2019.
[9] J. Zhao and J. Guo, "Big data analysis technology application in agricultural intelligence decision system," *IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, 2018, pp. 209-212, doi: 10.1109/ICCCBDA.2018.8386513.
[10] K. Jha, A. Doshi, P. Patel, and M. Shah, "A comprehensive review on automation in agriculture using artificial intelligence," *Artificial Intelligence in Agriculture*, pp. 1-12, 2019, doi: 10.1016/j.aiia.2019.05.004.
[11] B. Maurya, M. R. Beg, and S. Mukherjee, "Expert system design and architecture for farming sector," *2013 IEEE Conference on Information & Communication Technologies*, 2013, pp. 10-15, doi: 10.1109/CICT.2013.6558053.
[12] S. Pouyanfar *et al*., "Survey on deep learning: algorithms, techniques and applications," *ACM Computing Surveys*, vol. 51, no. 5, pp. 1-36, doi: 10.1145/3234150.
[13] J. Yadav, S. Chopra, and M. Vijayalakshmi, "Soil analysis and crop fertility prediction using machine learning," *International Journal of Innovative Research in Advanced Engineering*, vol. 8, no. 3, pp. 41-49, doi: 10.26562/ijirae.20210.v0803.003.
[14] J. Trontelj and O. Chambers, "Machine Learning Strategy for Soil Nutrients Prediction Using Spectroscopic Method," *Sensors*, vol. 21, no. 12, 2021, doi: 10.3390/ s21124208.
[15] L. Benos *et al*., "Machine learning in agriculture: a comprehensive updated review," *Sensors (Basel)*, vol. 21, no. 11, 2021, doi: 10.3390/s21113758.
[16] N. K. Manaswi, "Understanding and Working with Keras. In: Deep Learning with Applications Using Python," *Apress, Berkeley*, pp. 31-43, 2018, doi: 10.1007/978-1-4842-3516-4_2.
[17] A. Suruliandi, G. Mariammal, and S. P. Raja, "Crop prediction based on soil and environmental characteristics using feature selection techniques," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 27, no. 1, 2021, pp. 117-140, doi: 10.1080/13873954.2021.1882505.

[18] S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika, and J. Nisha, "Crop recommendation system for precision agriculture," *2016 Eighth International Conference on Advanced Computing (ICoAC)*, 2017, pp. 32-36, doi: 10.1109/ICoAC.2017.7951740.
[19] A. K. Mariappan, C. Madhumitha, P. Nishitha, and S. Nivedhitha. "Crop Recommendation System through Soil Analysis Using Classification in Machine Learning," *International Journal of Advanced Science and Technology*, no. 29, no. 3, pp. 12738-12747, 2020.
[20] N. Saranya and A. Mythili, "Classification of soil and crop suggestion using machine learning techniques," *International Journal of Engineering Research and Technology (IJERT)*, vol. 9, no. 2, 2020, doi: 10.17577/IJERTV9IS020315.
[21] N. Mahendra, "Crop predictions and machine learning approaches," *IJERT*, vol. 8, no. 8, 2020. doi: 10.17577/IJERTV9IS080029.
[22] R. Chaudhari, S. Chaudhari, A. Shaikh, R. Chiloba, and T. D. Khadtare "Fertility Prediction Using Data Mining Techniques," *International Journal of Future Generation Communication and Networking*, vol. 13, no. 3, pp. 213–220, 2020.
[23] P. Bandara, T. Weerasooriya, T. H. Ruchirawya, W J M Nanayakkara, M. A. C. Dimantha, and M. G. P. Pabasara, "Crop Recommendation System," *International Journal of Computer Applications*, vol. 175, no. 22, pp. 22-25, 2020, doi: 10.5120/ijca2020920723.
[24] G. D. Diniz *et al*., "Simplified tea classification on reduced composition profile using LDA," *Journal of Food Composition and Analysis*, vol. 39, 2015, doi: 10.1016/j.jfca.2014.11.012.
[25] A. S. Fotheringham and L. G. Reeds, "An Application of Discriminant Analysis to Agricultural Land Use Prediction," *Economic Geography*, vol. 55, no. 2, pp. 114–22, 1979, doi: 10.2307/142654.

BIOGRAPHIES OF AUTHORS

Varshitha D. N. is a research scholar at Sir M Visvesvaraya Institute of Technology, Bengaluru, Karnataka, India. She has 5 years of teaching and research experience. She holds a Master degree in Computer science & Engineering from VTU PG centre, Mysuru and Bachelor degree in Information Science & Engineering from NIEIT, Mysuru, Karnataka, India. Her research interests include datamining, machine learning, deep learning & DBMS. She can be contacted at email: varsha.swara@gmail.com.

Dr. Savita Choudhary is an associate professor working at Sir M Visvesvaraya Institute of Technology, Bengaluru, Karnataka. She has been awarded PhD for her research work in the area of soft computing from Bhagawanth university, Rajasthan. She has done Masters in Software Engineering from Banasthali university, Rajasthan, and Bachelor of Engineering from university of Rajasthan, India. She has 12+ years of teaching and research experience, she has guided several UG students and currently guiding four PhD students. She has published and presented papers in many international journals and conferences. Her research interests include soft computing, machine learning, and IoT. She can be contacted at email: choudhary7.mvit@gmail.com.