

**How to Cite:**

Suma, T., Kumar, S. V., & Sandhya, B. R. (2022). Classification of soil and prediction of fertilizer for specific crop cultivation using machine learning technique. *International Journal of Health Sciences*, 6(S8), 5468–5473. Retrieved from <https://sciencescholar.us/journal/index.php/ijhs/article/view/13475>

## **Classification of soil and prediction of fertilizer for specific crop cultivation using machine learning technique**

**Dr T. Suma**

Associate Professor, Computer Science and Engineering, Sri Venkateswara College of Engineering, Bangalore, India  
Corresponding author email: [tsuma.cs@gmail.com](mailto:tsuma.cs@gmail.com)

**Mrs Sathya Vijay Kumar**

Assistant Professor, Computer Science and Engineering, Sri Venkateswara College of Engineering, Bangalore  
Email: [sathyaviji80@gmail.com](mailto:sathyaviji80@gmail.com)

**Mrs Sandhya B R**

Assistant Professor, Information Science and Engineering, Sri Venkateswara College of Engineering, Bangalore  
Email: [sandhyabr20@gmail.com](mailto:sandhyabr20@gmail.com)

**Abstract**---Since India is an agricultural country, its economy is centered on agricultural yield growth and agroindustry products. In agriculture, yield prediction is a major challenge. All farmers want to know what kind of yields they can expect. In the past, yield forecasts were made taking into account the farmer's experience in specific fields and crops. Yield forecasting is a big problem and needs to be resolved based on the available data. Machine learning technology is suitable for this. To evaluate and forecast future crop yield, agriculture employs a range of machine learning approaches. A variety of machine learning techniques are used and evaluated in agriculture to estimate next year's yields. This paper proposes and implements a system for predicting yields from historical data. This is performed by examining agricultural data and utilising machine learning techniques such as support vector machines and random forests to determine the appropriate fertiliser for each crop. Analyze the many associated factors such as location and pH level, which are used to determine the alkalinity of the soil. Additionally, percentage of vitamins such as Nitrogen (N), Phosphorous (P), and Potassium (K) are determined using various packages such as APIs for climate and temperature, soil type, nutrient fee of the soil in that region, rainfall quantity within the region, and soil composition. All of the information's properties will be

reanalyzed, and the data will be educated using a variety of machine learning methods to create a version. The system comes with a version that is particular and accurate in predicting crop yield and providing the give up person with proper pointers about required fertiliser ratio based on atmospheric and soil factors of the land, which serves to boost farmer revenue by increasing crop output.

**Keywords**---agroindustry, API, nutrients, fertilizer, pH value.

## **Introduction**

Data evaluation is a technique for cleansing and modeling data in order to uncover important facts and conclusions.. It is a technique for analyzing, extracting, and predicting important facts from a vast set of data in order to extract a few patterns. Companies use this method to convert their customers' data into usable information. This study can potentially be used in the agricultural industry. Most farmers rely on their long-term experiences in the field on specific crops to predict a higher yield in the coming harvesting season, yet they still don't get a fair price for their commodities.

It usually occurs as a result of incorrect irrigation or poor crop selection, although it can also occur when the crop output is lower than projected. Agricultural researchers established the need for an effective technique to anticipate and improve crop growth, and the majority of agricultural research focuses on biological mechanisms to detect crop growth and increase production. Crop production is influenced by factors such as crop style, seed type, and environmental factors such as sunshine (temperature), soil (ph), water (ph), rainfall, and humidity. We can anticipate the simplest crop to own higher crop output by assessing the soil and atmosphere in a specific place. This forecast will help farmers choose the best crops for their farm based on soil type, temperature, humidity, water level, spacing depth, soil PH, season, and fertilizer use. Crop yield estimation can be challenging due to a variety of factors such as crop cultivar genetic potential, soil, weather, and cultivation practices (date of showing, amount of irrigation and fertilizer, etc.) as well as biotic stress. Crop yield estimation methods such as statistical, agro-meteorological, empirical, biophysical, and mechanistic have all been developed. India is a densely populated country, and the world's food supply must be ensured in the face of uncertain climate change. Suggestions for fertilizer use may assist farmers in determining the optimal fertilizer choice for their farming situation based on environmental factors.

In a number of research, information and communication technology (ICT) has been shown to be useful in calculating agricultural productivity. Since ancient times, agriculture has been the backbone of our country. Climate change is now a common occurrence. As a result of this lack of understanding of weather conditions, cultivating crops is difficult. We need to employ technology to identify or grasp crop data so that we can advise farmers on the best ways to grow crops. Fertilizer is also essential for successful crop growth.If fertilizer is applied in excess or insufficiently in the field, the soil's fertility may be lost, and the crop

may fail to produce the intended yield. As a result, fertilizer plays a significant role. Understanding temperature conditions is very important for India since crop forecasting can help to strengthen the Indian economy. The goal is to create an efficient model that can estimate crop production and provide fertilizer recommendations for a specific *crop*.

### **Related work**

Agricultural production is controlled by perception, average, minimum, and maximum, according to Eswari et al 2018 [1]. They have included a new metric called crop evapotranspiration in addition to temperature. The weather and the plant's growth stage have an impact on crop evapotranspiration. When calculating the yield of the groups, this element is taken into account. They all gathered data with these qualities and sent it to the Bayesian network, which divided it into two categories: true and false, and used a confusion matrix to compare it to the model's observed classifications to determine accuracy. Finally, they discovered that utilising Nave Bayes and Bayesian networks to predict agricultural yields is more accurate than using the SMO classifier, and that forecasting crop production in various climatic and farming conditions will be advantageous.

Historical climate and crop production data can be extracted using data mining approaches, according to Shruti Mishra et al 2018 [2]. It is possible to generate multiple crop yield projections. To allow farmers to make educated decisions about the soil and crop to be grown, a decision support system must be developed. They gathered the information, which included crop season, area, and hectare production, and examined it in WEKA with multiple algorithms. They assessed data using four different approaches, determining their correctness and comparing them to one another. J48, IBK, LAD tree, and LWL are the four approaches employed in WEKA. They came to the conclusion that the IBK was more accurate than others, albeit this was dependent on the dataset's type and character.

### **Existing Method**

#### **Handling in The Laboratory**

When the samples arrive at the soil testing laboratory, they should be subjected to several tests for various criteria as quickly as possible. If a fresh dirt is brought in that isn't listed in the records, the process will take longer.

#### **Drying of Samples**

When samples are received in the laboratory, they may lose moisture. Each sample's identification should be preserved during the preparation process.

#### **Post Drying Care**

The samples are air dried to make soil handling easier, but the soil is deteriorated and some of its organic constituents may be lost in the process.

A laboratory number is assigned to each sample, and both the field label and the laboratory number are recorded in a logbook. An undisturbed sample is extracted from the bulk sample if bulk density or other fabric-related analyses are to be performed. Undisturbed samples are frequently collected separately and not "processed" for further analysis. The samples are taken out of their containers and laid on white butcher paper (if desired) or in trays in an open area to dry. The sample is mixed every day (or several times per day in hot conditions) to ensure speedy and even drying. Clay-rich samples are examined on a regular basis to ensure that they can be crushed at the proper moisture content. The proper moisture content is below the liquid limit but above the rock hardness of the sample. Each sample requires a different moisture level, and determining the right one is a question of trial and error.

When the sample is dry, it is gently crushed with a wooden mallet or equivalent tool (if possible). The procedure's purpose is to shatter soil aggregates without fracturing any rock, gravel, or individual mineral grains. This is especially challenging in soil samples with substantial amounts of soft or brittle gravel or rock fragments. Until the materials are separated, the entire sample is sieved, crushed, and re-sieved. If a large number of samples are needed, the field sample can be split and a portion crushed and sieved. As needed, stored samples are subsampled for subsequent processing.

#### **Disadvantages of existing system**

1. It will take 15-20 days to receive the results.
2. The overall efficiency of the land cannot be predicted by a single soil sample.
3. The farmer lacks the necessary information to determine which fertilizer to apply.
4. Chemical procedures used in the laboratory can be hazardous at times, thus caution is advised.
5. Farmers are unable to transport many soil samples to the lab.

#### **Proposed system and implementation**

Our solution will overcome the limitations of the current crude methods. The sensors that we employed here are measuring soil factors like as pH, moisture, and humidity. Using machine learning, provide a farmer with a crop yield depending on land acreage, rainfall, temperature, and district. The collected values are then transferred to our server via wi-fi, where they are compared to the datasets available. Machine learning algorithms are used to determine which crop is most suited for that soil sample, as well as which fertilisers should be used and the expected yield cost. The result is shown on the graphical user interface (GUI) that we will be developing.

The methods being used are:

- Data collection
- Data pre-processing
- Training and Testing
- Applying algorithm

- Prediction and Accuracy

Data collection: Prior to releasing a machine learning model into production, it is critical that an organisation has the capacity to train, test, and validate it. Datasets are collected in this module from a variety of sources, including GKVK, DES, agmarknet.gov.in, and krishimaratavahini, as well as websites like Kaggle, UCI, and Data towards science. The databases include information from the preceding five years on Ph, soil moisture, price, area, production, and yield. Data is gathered for a number of districts.

Data pre-processing : Unwanted properties from our datasets are removed during data pre-processing. Feature extraction is used to extract only the attributes that affect a crop's price and yield, such as location, acreage, and variety. , as well as output and yield. Importing the data set, dealing with missing data, encoding categorical data, and so on. Feature scaling and splitting the data set into training and test sets.

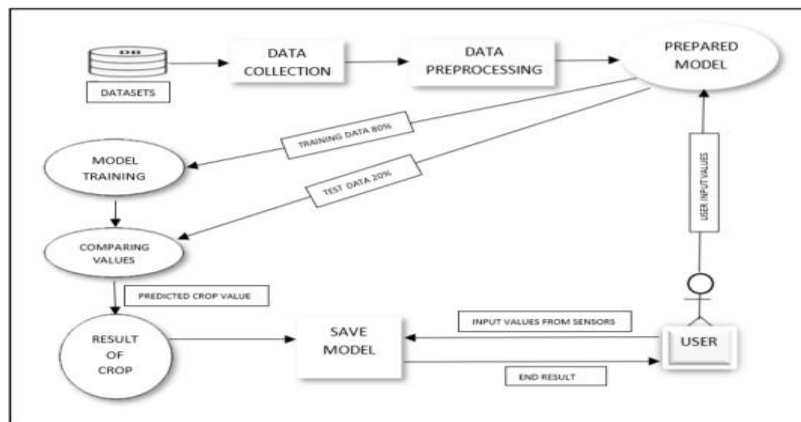
Training and Testing : In order to provide complex findings, the training data is an initial set of data needed to help determine which crop is suitable for the given soil properties Testing: The training dataset in our research is made up of input values for several soil properties.

Applying Algorithm: A Support Vector Machine (SVM) is a discriminative classifier whose formal definition is a separating hyperplane. To put it another way, given labelled training data gleaned from several crop prediction databases, the method generates an ideal hyperplane that identifies the best crop for the given soil.

## **System design**

### **Data flow diagram**

A dataflow diagram is a graphical representation of data "flow" across an information system that simulates process characteristics. A DFD is typically used as a first step to give a high-level overview of the system without going into considerable detail, which can then be elaborated upon later. Data processing can also be visualized using DFDs. A DFD shows the many types of data that will be input to and output from the system, as well as how the data will flow through the system and where it will be kept.



## Conclusion

Based on the processing of numerous soil factors such as pH, moisture, and humidity, soil analysis and crop prediction provides crop and fertilizer recommendations. Soil testing services can be delivered to farmers' homes. This concept effectively replaces the traditional method of soil testing, allowing farmers to learn more about their soil in a more convenient and timely manner. This idea's outcome assists farmers in making decisions on what to cultivate and prevents them from applying imbalanced fertilisers. This paper is also very helpful in estimating the cost of their products

## References

1. A.T.M Shakil Ahamed, Navid Tanzeem Mahmood, Nazmul Hossain, Mohammad Tanzir Kabir, Kallal Das, Faridur Rahman, Rashedur M Rahman "Applying Data Mining Techniques to Predict Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in Bangladesh" 978-1-4799-8676-7/15/\$31.00 copyright 2015 IEEE SNPD 2015, June 1-3 2015
2. Crop Yield Prediction and Efficient use of Fertilizers S.Bhanumathi, M.Vineeth and N.Rohit. International Conference on Communication and Signal Processing, April 4-6, 2019, India.
3. Dakshayini Patil et al, "Rice Crop Yield Prediction using Data Mining Techniques: An Overview", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 7, Issue 5, May 2017.
4. K.E. Eswari. L.Vinitha. " Crop Yield Prediction in Tamil Nadu Using Bayesian Network ", International Journal of Intellectual Advancements and Research in Engineering Computations, Volume-6, Issue 2, ISSN: 2348-2079.
5. Shruti Mishra, Priyanka Paygude, Snehal Chaudhary, Sonali Idate "Use of Data Mining in Crop Yield Prediction" IEEE Xplore Compliant - Part Number: CFP18J06-ART, ISBN: 978-1-5386-0807-4; DVD Part Number: CFP18J06DVD, ISBN: 978-1-5386-0806-7.
6. Sorghum Yield Prediction using Machine Learning Judicael Geraud N. ZANNOU University of Abomey Calavi (UAC) Abomey-Calavi, Benin.