

“Crop Yield Prediction and Soil Evaluation Using Machine Learning”

*Note: Sub-titles are not captured in Xplore and should not be used

Mrs. Priyanka Shinde
Electronics and Telecommunication
Engineering
JSPM's Rajarshi Shahu College Of
Engineering
Pune, India
priyanka.ns02@gmail.com

Sumit Mukhedkar
Electronics and Telecommunication
Engineering
JSPM's Rajarshi Shahu College Of
Engineering
Pune, India
sumitmukhedkar2001@gmail.com

Venugopal Birla
Electronics and Telecommunication
Engineering
JSPM's Rajarshi Shahu College Of
Engineering
Pune, India
ksbirla2025@gmail.com

Akash Pandule
Electronics and Telecommunication
Engineering
JSPM's Rajarshi Shahu College Of
Engineering
Pune, India
akashpandule2000@gmail.com

Aditi Gosavi
Electronics and Telecommunication
Engineering
JSPM's Rajarshi Shahu College Of
Engineering
Pune, India
aditigosavi1411@gmail.com

Abstract—Agriculture is very important part of India's economy and helps ensure there is enough food for everyone. However, climate change is causing unpredictable weather patterns, which can make it harder for farmers to grow crops successfully. To help farmers choose which crops to plant, a computer program is being developed using a technology called "machine learning". This program uses a method called "Naive Bayes" to predict which crops are likely to grow well based on factors like temperature, humidity, and moisture. An Web application is being created, which will ask farmers to input data about their soil type and get predictions about which crops to plant. This will make it easier for farmers to make informed decisions about what to grow, and hopefully, help them to have a more successful harvest.

Keywords— Agriculture, Machine Learning, Native Bayes, Crop prediction .

I. INTRODUCTION (HEADING 1)

From ancient days, agriculture is considered as the main source of supply to satisfy the daily needs of human lives. It is also considered a primary occupation, and also one of the India's major industrial sectors. The farmers are ought to follow a traditional naked eye observation and yielded healthy crops without the involvement of chemicals for animals and also to their cultivation land in order to keep healthy diversity. But nowadays, weather conditions are being rapidly changing against the elemental assets to deplete the food and increase the security. In meantime, the GDP in agricultural sector is

keep on decreasing, where in 2005 it was about 17.211.1, in 2018 it was 52020 it came down to 2 frames come from rural areas, and if the revenue from crop production goes down, their lifestyle would be influenced by the farms at industry level. Agriculture is having a great impact on the country's economy. In the last decade India has seen serious natural calamities like drought or flood. Due to such disasters, there is a huge loss to crop production and ultimately to the farmers. Due to such financial loss many farmers are committing suicide. If natural calamities are not present then there may be sudden pest attack destroying the crop. In any case farmer and the crop are always at the edge of risk. Production of crops depends on four main factors like climate, soil fertility, availability of water, and disease or pests. And four biological factors as organic matter content, Activation carbon content, Nitrogen content, and root health. The health of soil can be tested in the range of 1 to 100[1]. A soil health test report provides an integrative assessment and also identifies specific soil constraints.

In this project we can overcome all these problems using machine learning approach. The production of crops may depend on geographical conditions of the region like river ground, hill areas or the depth areas. Weather conditions like humidity, rainfall, temperature, cloud. Soil type may be clay, sandy, saline or peaty. Soil composition can be copper, potassium, phosphate, nitrogen, manganese, iron, calcium, pH value or carbon and different methods of harvesting. Many parameters are used for different crops to do different predictions. These prediction models can be studied by using researches. These predictions are classified as two types. One is traditional statistic method and other is machine learning techniques. Traditional method helps in predicting single sample spaces. And machine learning methods help in predicting multiple predictions. We need not to consider the structure of data models in traditional method whereas we need to consider the structure of data models in machine learning methods. Here the main functionality is to predict the crop yield using the image given by the user. In this project we are taking input as image of the soil then Dataset as image will preprocess the image of soil and it will clean dataset and reduced noise data. Then feature extraction will be done

extract all the features. Then Segmentation will be done and image is divided into multiple parts after that Train model split the database into Testing 20% and Training 80%. It will classify using CNN (Convolutional Neural Network) algorithms. Output will be predicting the soil type and predict the suitable crop.

II. LITERATURE SURVEY

In reference [1] Prof. A. V. Deorankar et al. Have concluded that this paper helps in improving the yield rate of crops. The author said that Production of crops depends on four main factors like climate, soil fertility, availability of water, and disease or pests. And four biological factors as organic matter content, Activation carbon content, Nitrogen content, and root health the growth of plants are depend on internal and external factor, the external factors are water, light, nutrients, and temperature and the internal factors are nitrogen (N), phosphorus (P) and potassium (K). It will tell which nutrients is less in the soil and ask to add that particular nutrient.

In reference [2] Ramesh Medar, V. S. Rajpurohit and S. Shweta at al. In this paper author has concluded that they are using two different methods first is K nearest neighbor and native bias method. Here they have created java application which includes three parts first is managing datasets second is testing dataset and third is analyzing dataset. They are taking previous year crop details e.g. crop= rice, district= Pune, season= Kharif, yield= good this is the example of dataset by using this dataset they are predicting the crop yield. They are using weka tool in project all the datasets are converted attribute relation in format.

In reference[3] Yogesh Gange and Sandhya In this paper author has concluded that they are using different data set soil prediction. Soil Prediction needs different kinds of data gathered from different sources like meteorological data, agri-meteorological, soil(pH,N,P,K) data, remotely sensed data, agricultural statistics. To handle such a huge data the best option we have is Data Mining. Data Mining and using different data set to different soil prediction like Soybin they are using Decision tree analysis data set and rice they are using multiple linear Regression data and this data set and accuracy is 90 to 95 percentage. They are taking different data set to different soil prediction and it accuracy rate is above 85% suppose using super vector data set it use any crop prediction . The results show that support vector regression can serve as a better reference model for yield prediction. It is computationally less demanding.

In reference [4] P. S. Vijayabaskar, Sreemathi. R, Keertanaa. E In this paper author has conclude that they are using two different algorithms for two different perspective i.e. CNN Convolutional neural network for detecting wild animal intrusion and SVC support vector machine for detecting soil parameters and predict the crop by which it helps to reduce loss of money for growing unwanted crop. In CNN an USB camera is used to capture the image. So by denting the image wild animal such as lion when it comes to entry level it will send a sms or email to respected farmer.

In reference [5] Machine learning (ML) has emerged as a valuable tool for addressing the crop yield issue. By utilizing supervised learning, ML algorithms can predict a target outcome based on a given set of predictors. Crop yield prediction involves analyzing past historical data, including factors such as temperature, humidity, pH, rainfall, and crop

name. This helps farmers to determine which crop will yield the best results under current weather conditions. One commonly used ML technique for crop yield prediction is the Random Forest algorithm, which provides accurate crop predictions by considering the least number of models. By accurately predicting crop yields, farmers can make more informed decisions about planting and harvesting, leading to better crop yields and more efficient use of resources in the agriculture sector.

In reference [6] S. V. Bhosale, R. A. Thombare, P. G. Dhemey and A. N. Chaudhari at al. The agricultural sector generates vast amounts of data through various sources, making it a big data industry. Smart technologies are now utilized to collect data through electronic devices. To optimize crop yield, agricultural data must be analyzed and mined using technologies such as data analytics and machine learning. In one particular project, the agricultural data will be analyzed using these technologies to produce useful results that can be shared with farmers to improve efficiency and productivity. By providing farmers with this information, they can make informed decisions to improve crop yield and reduce waste in the agricultural sector.

III. METHODOLOGY

Using a Convolutional Neural Network (CNN) for crop yield prediction and soil evaluation involves a different approach than traditional ML techniques. CNNs are commonly used in image recognition tasks but have also been applied to other fields such as agriculture.

The methodology for using CNN for crop yield prediction and soil evaluation involves the following steps:

1. Data Collection: Collect a large dataset of agricultural data that includes images of crops and soil samples, as well as other factors such as temperature, humidity, pH, and rainfall.

2. Preprocessing: The collected images need to be preprocessed before feeding them into the CNN. This includes resizing, normalization, and image augmentation to increase the size of the dataset.

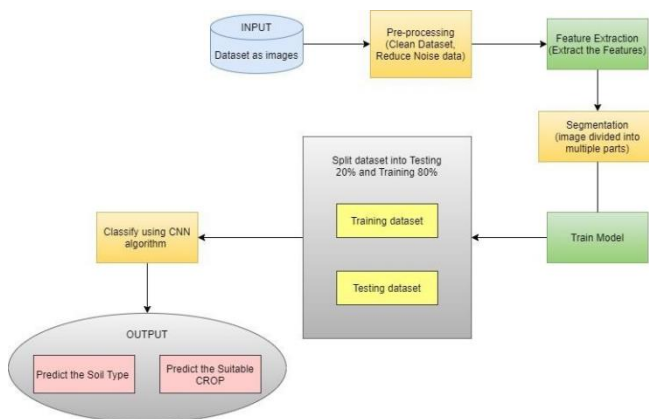
3. Training: The CNN is trained using the preprocessed dataset. The training process involves adjusting the weights of the CNN layers to minimize the error between the predicted and actual crop yields and soil properties.

4. Testing: The trained CNN is tested on a separate dataset to evaluate its performance in predicting crop yields and soil properties. The accuracy of the CNN is measured using metrics such as Mean Absolute Error(MAE) and Root Mean Squared Error(RMSE).

5. Analysis: The results of the testing phase are analyzed to identify the most important factors in predicting crop yields and soil properties. The analysis can also reveal the impact of different environmental factors such as temperature, humidity, and rainfall on crop yields and soil properties.

The use of CNN for crop yield prediction and soil evaluation offers several advantages. Firstly, it can provide more accurate predictions than traditional ML techniques as it can detect patterns in images that may not be visible to the human eye. Secondly, it can be used to analyze large amounts of data quickly and efficiently, enabling farmers to make informed decisions about planting and harvesting. Finally, it

can help in the development of more sustainable agricultural practices, reducing the environmental impact of farming.



A Brief overview of the Crop Yield Prediction System:

1. In the development of a crop prediction system, the selection of the agriculture field and crop of choice is the first step.
2. The input data needed for the system includes information about the soil, such as Nitrogen (N), Phosphorus (P), Potassium (K) content, and micronutrients present in the soil, and moisture levels collected over a period of time.
3. Data preprocessing is necessary to eliminate redundant and inconsistent data.
4. Important features should be extracted through attribute selection.
5. An appropriate and efficient classification algorithm should be employed for the system.
6. The prediction or recommendation results obtained can be provided to the farmers for decision-making purposes.

IV. FUTURE SCOPE

The research in the field of crop yield prediction and soil evaluation using machine learning faces several challenges. In order to overcome these challenges, further study and exploration are required. To improve the performance of the prediction system, a more robust and novel classifier should be employed. The research should also focus on recommending the most suitable crop for a farmer's land based on the nutrient content of the soil. Additionally, it is important to suggest water-tolerant seed varieties for sowing so that in case of flood or drought, the crop can withstand the damage. The use of sensors to measure soil moisture and nutrient levels can provide valuable information to guide the farmers in their agricultural practices.

CONCLUSION

The agriculture sector is crucial for the economic growth of a country, but it has been slow in adopting new technologies

Identify applicable funding agency here. If none, delete this text box.

such as machine learning. To improve the yield rate of crops and solve agriculture-related problems, farmers need to be aware of these new techniques. Several machine learning techniques have been applied to agriculture to enhance crop yield and accuracy. The accuracy of yield can be improved by comparing different methods and performance metrics. Accurate crop yield prediction is vital for decision-makers at national and regional levels to make informed decisions. Various approaches can be used for crop yield prediction.

REFERENCES

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

- [1] A. V. Deorankar and A. A. Rohankar, "An Analytical Approach for Soil and Land Classification System using Image Processing," 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2020, pp. 1416-1420, doi: 10.1109/ICCES48766.2020.9137952.
- [2] R. Medar, V. S. Rajpurohit and S. Shweta, "Crop Yield Prediction using Machine Learning Techniques," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 2019, pp. 1-5, doi: 10.1109/I2CT45611.2019.9033611.
- [3] Y. Gandge and Sandhya, "A study on various data mining techniques for crop yield prediction," 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICECCOT), Mysuru, India, 2017, pp. 420-423, doi: 10.1109/ICECCOT.2017.8284541.
- [4] P. S. Vijayabaskar, R. Sreemathi and E. Keertanaa, "Crop prediction using predictive analytics," 2017 International Conference on Computation of Power, Energy Information and Communication (ICCPEIC), Melmaruvathur, India, 2017, pp. 370-373, doi: 10.1109/ICCPEIC.2017.8290395.
- [5] Y. J. N. Kumar, V. Spandana, V. S. Vaishnavi, K. Neha and V. G. R. R. Devi, "Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector," 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2020, pp. 736-741, doi: 10.1109/ICCES48766.2020.9137868.
- [6] S. V. Bhosale, R. A. Thombare, P. G. Dhemey and A. N. Chaudhari, "Crop Yield Prediction Using Data Analytics and Hybrid Approach," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697806.
- [7] R. J.V.K.G. Kalaiselvi, A. Sheela, D. S. D. and J. G., "Crop Yield Prediction Using Machine Learning Algorithm," 2021 4th International Conference on Computing and Communications Technologies (ICCCT), 2021, pp. 611-616, doi: 10.1109/ICCCT53315.2021.9711853.
- [8] R. Nikhil, B. S. Anisha and R. Kumar P., "Real-Time Monitoring of Agricultural Land with Crop Prediction and Animal Intrusion Prevention using Internet of Things and Machine Learning at Edge,"

2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2020, pp. 1-6, doi: 10.1109/CONECCT50063.2020.9198508.

- [9] J. E. Sánchez-Galán, J. S. Reyes, J. U. Jiménez, E. I. Quirós-McIntire and J. R. Fábrega, "Supervised Classification of Spectral Signatures from Agricultural Land-Cover in Panama Using the Spectral Angle Mapper Algorithm," 2019 XLV Latin American Computing

Conference (CLEI), Panama, Panama, 2019, pp. 1-7, doi: 10.1109/CLEI47609.2019.235101.

- [10] D Ramesh , B Vishnu Vardhan, Analysis Of Crop Yield Prediction Using Data Mining Techniques, International Journal of Research in Engineering and Technology, Jan-2015.