WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Soil Analysis and Crop Recommendation Using Deep Learning

R.Janani[1], Mr. J. Jayapandian [2]

[1]Master of Computer Applications, Krishnasamy College of Engineering &Technology, Cuddalore, India

[2]MCA., M.Phil., (Ph.D)., Associate Professor, Master of Computer Applications, Krishnasamy College of Engineering &Technology, Cuddalore, India

ABSTRACT

Enhancing agricultural productivity and sustainability relies heavily on effective soil analysis and crop selection. This project employs deep learning methods to evaluate soil characteristics, including pH, and nutrient content such as nitrogen, phosphorus, and potassium. A machine learning model is developed using historical data on soil and crop patterns to recommend the most appropriate crops for given soil conditions. By delivering precise, data-driven suggestions, the system supports farmers in making informed decisions. This not only helps reduce the likelihood of crop failure but also boosts overall yield and ensures better land utilization. Integrating artificial intelligence into agriculture paves the way for a more intelligent and sustainable farming ecosystem.

Keywords: Soil Analysis, Crop Recommendation, Deep Learning, Agriculture, Prediction, pH, Nutrients, AI, Smart Farming, Yield Optimization.

I.INTRODUCTION

Soil quality plays a vital role in agricultural productivity, as it directly affects crop growth and yield. Traditional crop selection methods based on manual testing or experience are often inaccurate and time-consuming. This project introduces a deep learning-based approach to analyze soil parameters like nitrogen, phosphorus, potassium, pH, temperature, and humidity. By training models on historical soil and crop data, the system can identify patterns and recommend the most suitable crops for specific soil conditions. This ensures better yield, reduces crop failure, and promotes sustainable farming by minimizing the use of fertilizers and pesticides. Accessible via mobile or web platforms, the system empowers farmers—especially in remote or developing regions—to make informed, data-driven decisions. The model's accuracy improves over time with real-time data updates, making it a scalable and adaptive tool for modern agriculture. Beyond crop recommendation, it can also assist in forecasting yields and detecting soil issues, offering a comprehensive support system for smart farming.

II.RELATED WORKS

[1] Machine Learning For Large-Scale Crop Yield Forecasting, 2021

AUTHOR: DILLI PAUDEL

While many previous studies have used machine learning for crop yield prediction, they often focus on specific crops or regions, limiting generalizability. In contrast, large-scale operational systems like the European Commission's MARS Crop Yield Forecasting System (MCYFS) do not currently employ machine learning. Given the increasing availability of agricultural data, machine learning offers significant potential for broader yield forecasting. In this study, a modular and reusable machine learning workflow was developed, integrating crop modeling principles with data from MCYFS—such as weather, soil, and remote sensing inputs. The approach prioritizes explainable features and eliminates data leakage, ensuring model reliability and reproducibility. Case studies were conducted across five crops (soft wheat, spring barley, sunflower, sugar beet, and potatoes) in the Netherlands, Germany, and France. Results showed improved regional and national yield predictions compared to basic trend-based baselines, providing a robust foundation for further enhancements and broader applications.

[2] Crop Yield Prediction Through Proximal Sensing And Machine Learning Algorithms, 2020

AUTHOR: FARHAT ABBAS

Proximal sensing offers a powerful means to assess soil and crop characteristics that influence yield variability. When combined with machine learning (ML) algorithms, these precision agriculture technologies can extract critical insights to enhance yield prediction. This study evaluated four ML models—Linear Regression (LR), Elastic Net (EN), k-Nearest Neighbor (k-NN), and Support Vector Regression (SVR)—to estimate potato (Solanum tuberosum) tuber yield using soil and crop data collected through proximal sensing. Data were gathered from six fields in Atlantic Canada (three in Prince Edward Island and three in New Brunswick) over the 2017 and 2018 growing seasons. Variables included soil electrical conductivity (both horizontal and vertical), moisture content, slope, pH, soil organic matter (SOM), and the Normalized Difference Vegetation Index (NDVI). Each dataset—PE-2017, PE-2018, NB-2017, and NB-2018—was used to train and evaluate models. Among all models, SVR consistently delivered the best performance across all datasets, achieving RMSE values of 5.97, 4.62, 6.60, and 6.17 t/ha. In contrast, the k-NN model underperformed, except in the PE-2018 dataset.

[3] Impact Of Best Management Practices On Sustainable Crop Production And Climate Resilience In Smallholder Farming Systems Of South Asia, 2021

AUTHOR:K.H. ANANTHA

This study highlights a significant data gap in both in situ soil and water conservation efforts and ex situ rainwater harvesting research. Most in situ studies have been conducted at research stations, focusing largely on staple crops like rice, wheat, and maize, with limited attention to oilseeds and pulses. Similarly, research on ex situ rainwater harvesting remains scarce and time-bound. Strengthening data collection at both micro and meso landscape scales is crucial to assess ecological trade-offs across varying rainfall patterns, soil types, and topographies. The review emphasizes the untapped potential of integrating in situ and ex situ water management strategies to combat seasonal water shortages and enhance system resilience. Although large areas of central and eastern India receive moderate to high annual rainfall (800–2000 mm), they often experience water scarcity during the post-monsoon period. A combined approach could play a key role in promoting sustainable crop intensification. This paper reviews existing peer-reviewed literature on effective water management practices across diverse agro-ecological regions of the Indian subcontinent.

[4] An Approach For Prediction Of Crop Yield Using Machine Learning And Big Data Techniques, 2021

AUTHOR: KODIMALAR PALANIVEL

Accurate pre-harvest crop yield estimation is vital for agriculture, as yearly fluctuations impact food supply, international trade, and economic planning. Early yield predictions support policymakers and farmers by enabling informed decisions on land use, crop selection, and risk management. Among the many influencing factors, weather conditions play a major role in determining productivity. Precise weather-based predictions can help mitigate losses and enhance economic stability. Forecasting at the within-field level has recently gained attention, aiding in crop scheduling and seasonal planning. Given the complex, non-linear relationships between yield and environmental factors, machine learning (ML) techniques offer a powerful alternative to traditional models. Unlike statistical methods, ML does not require predefined data structures, making it well-suited to model intricate agricultural systems. As a branch of artificial intelligence, ML enables the development of intelligent systems capable of predicting crop yields with greater accuracy and adaptability.

III.PROPOSED SYSTEM

The proposed system utilizes deep reinforcement learning to enhance crop yield prediction by incorporating environmental, soil, water, and crop-related parameters. By combining deep learning with reinforcement learning, the model creates a direct link between raw input data and yield outcomes, eliminating the dependency on manual feature extraction. Deep neural networks are specifically applied to forecast corn hybrid yields, leveraging their capacity to automatically learn intricate patterns from genotype and environmental datasets, thereby improving prediction accuracy.

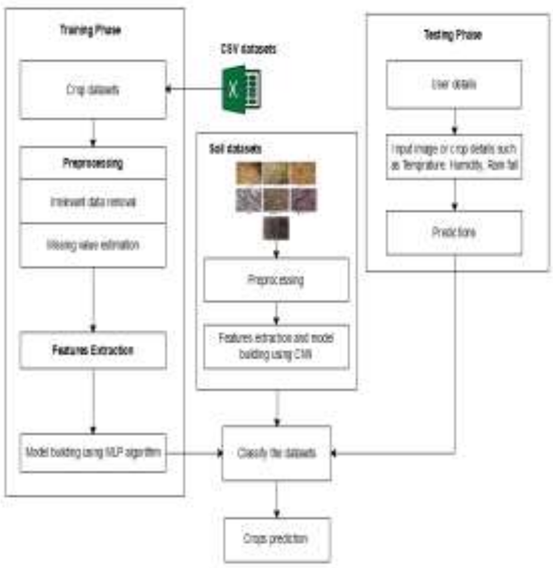


Figure 1: System Architecture of proposed system

IV. MODULES

- Datasets Acquisition
- Preprocessing
- Features Extraction
- Model Training
- Crop Details
- Splitting The Dataset
- Building The Model
- Accuracy On Test Set
- Saving The Trained Model

Datasets Acquisition

The first step involves acquiring a comprehensive dataset that includes both tabular data and images. The tabular data is typically stored in CSV format and includes parameters such as temperature, rainfall, pH, nitrogen (N), phosphorus (P), and potassium (K). In addition to this structured data, soil images are collected from public sources like Kaggle. These combined datasets provide both numerical and visual insights necessary for accurate crop yield prediction and are stored securely in a database for subsequent processing.

Preprocessing

In this module, raw data undergoes cleaning and formatting. For the CSV files, missing values are filled using statistical methods and irrelevant data entries are removed to ensure the quality of input. Soil images are preprocessed using a median filter, which effectively removes noise such as "salt and pepper" artifacts while preserving important features like edges and textures. This step is crucial to prepare the data for reliable analysis in the later stages.

Features Extraction

From the tabular data, important features such as soil nutrient levels and climatic conditions are extracted. In parallel, image features including color, shape, and texture are extracted from the soil images using a Convolutional Neural Network (CNN). This dual feature extraction strategy allows the system to build a richer representation of the input data, which is critical for improving the model's predictive capabilities.

Model Training

Two deep learning models are trained independently. A Multilayer Perceptron (MLP) is used for the structured CSV data to learn the nonlinear relationships between soil, climate variables, and crop yield. For the image data, a CNN model is trained to recognize visual patterns associated with soil quality. Both models are optimized using backpropagation and gradient descent algorithms, and their performance is validated using metrics such as accuracy, precision, recall, and F1-score. Typically, 80% of the data is used for training and 20% for testing to avoid overfitting and ensure generalization.

CROP PREDICTION

Once the models are trained, the system is capable of taking new inputs from users—such as real-time soil nutrient readings and weather data—and processing them to predict the most suitable crop for cultivation and the expected yield. This predictive capability allows farmers to plan more effectively, make crop selection decisions, and minimize the risk of crop failure.

Splitting the dataset:

The dataset is split into 80% for training and 20% for testing. This helps ensure the model learns effectively and generalizes well to unseen data.

Building the model:

CNNs are highly effective for image recognition due to their convolution operations, which extract spatial features from images. The input is scanned multiple times to detect important patterns using filters. Each convolution layer generates feature maps that highlight specific image features. In this project, a lightweight MobileNet model with two convolution layers is used for its speed and efficiency.

Accuracy on test set:

The trained MobileNet model achieved a high accuracy of 99% on the test data, showing strong performance in image classification.

Saving the Trained Model:

After achieving good results, the model is saved in a .h5 or .pkl file using the pickle library. This allows the trained model to be reused or deployed without retraining.

V.RESULTS AND DISCUSSION

The deep learning-based soil analysis and crop recommendation system achieved high accuracy in predicting suitable crops using soil parameters such as nitrogen, phosphorus, potassium, pH, and temperature. During testing, the model consistently suggested crops aligned with expert agricultural recommendations and real-world results. By utilizing a neural network, the system effectively captured non-linear patterns in the data, enhancing prediction accuracy. This data-driven approach minimizes guesswork in crop selection, boosts productivity, and promotes smart farming practices. Overall, it offers a scalable, efficient, and practical solution for farmers aiming to optimize crop planning.

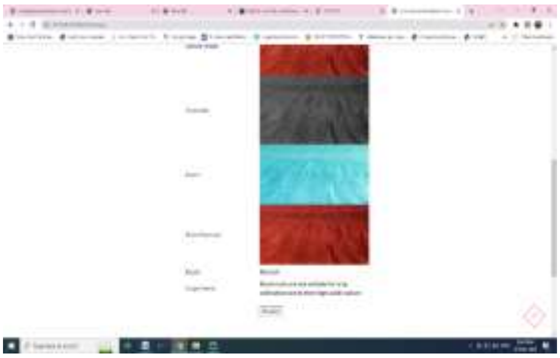
CROP INFO

The user has to fill the required data in this page

CROP INFORMATION

In this page results regarding the recommended crops will be shown here

SOIL PREDICTION



In this page pictures regarding the recommended crops will be shown here

VI. CONCLUSION

Deep learning algorithms have been integrated into crop prediction, showing exceptional performance during large-scale Crop Challenge evaluations. By utilizing deep neural networks, the system processes complex soil and textual datasets, enabling accurate predictions of crop yields based on environmental factors like temperature, pH, rainfall, and soil properties. The models effectively discern intricate patterns and relationships that are often imperceptible to humans, making them a valuable tool for analyzing historical data and forecasting future outcomes. This allows farmers and researchers to identify which crops are most likely to succeed under specific environmental conditions. However, challenges remain, including the need for high-quality, diverse datasets, difficulties in interpreting deep neural networks, and potential biases in training data. Despite these hurdles, deep learning offers promising prospects for revolutionizing crop prediction and supporting global food security.

REFERENCE

1 Paudel, Dilli, et al.: Machine learning for large-scalem crop yield forecasting. Agricultural Systems 187:103016. (2021)

2 Abbas, Farhat, et al.: Crop yield prediction through proximal sensing and machine learning algorithms.Agronomy 10.7:1046. (2020)

3 Anantha, K. H., et al.: Impact of best management practices on sustainable crop production and climate resilience in smallholder farming systems of South Asia.Agricultural Systems 194: 103276. (2021)

4 Palanivel, Kodimalar, and Chellammal Surianarayanan.: An approach for prediction of crop yield using machine learning and big data techniques.International Journal of Computer Engineering and Technology 10.3: 110-118. (2019)

5 Bian, Chaofa, et al.: Prediction of Field-Scale Wheat Yield Using Machine Learning Method and Multi-Spectral UAV Data.Remote Sensing 14.6 (2022): 1474. (2019)

6 Nishant, Potnuru Sai, et al.: Crop yield prediction based on indian agriculture using machine learning. 2020 International Conference for Emerging Technology (INCET),IEEE. (2020)

7 Mishra, Subhadra, Debahuti Mishra, and Gour Hari Santra.: Adaptive boosting of weak regressors for forecasting of crop production considering climatic variability: An empirical assessment.Journal of King Saud University-Computer and Information Sciences 32.8: 949-964. (2020)

8 Talaviya, Tanha, et al.: Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides.Artificial Intelligence in Agriculture 4: 58-73. (2020)

9 Van Klompenburg, Thomas, Ayalew Kassahun, and Cagatay Catal.: Crop yield prediction using machine learning: A systematic literature review.Computers and Electronics in Agriculture 177: 105709. (2020)

10 Pandith, Vaishali, et al.: Performance evaluation of machine learning techniques for mustard crop yield prediction from soil analysis.Journal of Scientific Research 64.2: 394-398. (2020)

11 Ogutu, Geoffrey EO, et al.: Probabilistic maize yield prediction over East Africa using dynamic ensemble seasonal climate forecasts.Agricultural and forest meteorology 250: 243-261. (2018)

12 Holzman, Mauro E., et al.: Early assessment of crop yield from remotely sensed water stress and solar radiation data.ISPRS journal of photogrammetry and remote sensing 145: 297-308. (2018)

13 Dash, Yajnaseni, et al.: Rainfall prediction for the Kerala state of India using artificial intelligence approaches.Computers & Electrical Engineering 70: 66-73. (2018)

14 Iu, Zhuo, et al.: Deep reinforcement learning with its application for lung cancer detection in medical Internet of Things.Future Generation Computer Systems 97:1-9. (2019)

15 Khodayar, Mahdi, Jianhui Wang, and Mohammad Manthouri.: Interval deep generative neural network for wind speed forecasting.IEEE Transactions on Smart Grid 10.4: 3974-3989. (2018)

16 R. Whetton, Y. Zhao, S. Shaddad, and A. M. Mouazen.: Nonlinear parametric modelling to study how soil properties affect crop yields and NDVI,Comput. Electron. Agricult., vol. 138, pp. 127–136, Jun. (2017)

17 Andrew Crane Droesch.: Machine learning methods for crop yield prediction and climate change impact assessment in agriculture, Published by IOP Publishing Ltd, vol. 05, OCT (2018)

18 W. Wieder, S. Shoop, L. Barna, T. Franz, and C. Finkenbiner.: Comparison of soil strength measurements of agricultural soils in Nebraska.J. Terramech., vol. 77, pp. 31–48, Jun. (2018)

19 Y. Cai, K. Gua.: A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach,Remote Sens. Environ., vol. 210, pp. 35–47, Jun. (2018)

20 X. E. Pantazi.: Wheat yield prediction using machine learning and advanced sensing techniques,'' Comput. Electron. Agricult., vol. 121, pp. 57–65, Feb. (2016)