RESEARCH ARTICLE OPEN ACCESS

Intelligent Solution for Crops Using Crop Recommendation Fertilizer Suggestion and Plant Disease Detection

[1]Nayeni Vamsi, [2]Mrs. Y. Nirmala
[1](Student, Dept. of Master of Computer Applications, Amrita Sai Institute of Science and Technology, Paritala, Andhra Pradesh, 521180, India.
Email: vamsichinnu208@gmail.com)
[2](Asst. Prof, Dept. of Computer Science & Engineering, Amrita Sai Institute of Science and Technology, Paritala, Andhra Pradesh, 521180, India.
Email: nirmala.y24@gmail.com)

--************************----------------------------------

Abstract:

An integrated agricultural decision support system is introduced in this report to help improve yield and the health of crops using data analysis. There are three main modules within the system: a crop recommendation tool, a suggestion for using fertiliser and detection of plant diseases. The system uses decision trees and random forests to pick out the best crops for farmers using data about N, P, K, humidity, temperature, pH and rainfall. The fertiliser suggestion module determines the best regimes for each crop and uses that information to fix fertility issues in soil and help crops develop well. The module uses CNNs to catch early indications of disease or pests on leaves caught digitally. Spotting the disease early helps you take action to reduce losses. This combination of systems gives a detailed solution for improving crop care, making better use of resources and dealing with agricultural risks. In the report, the method, execution, results and future goals for each module are analysed, emphasising the ways technology might transform agriculture.

Keywords *— Agricultural Decision Support System, Crop Recommendation, Fertilizer Suggestion, Plant Disease Detection.*

--************************----------------------------------

1. Introduction

Much of the Indian population works in agriculture and the agriculture sector gives its citizens enough food to eat. As more people are added to the world's population, more food is required which is making sure we have enough food difficult. A number of nations, like India, remain challenged by food shortages, caused in part by the rise in their population. By raising agricultural productivity, it becomes much easier to prevent hunger and famine. The United Nations has made food security and fewer lives suffering from hunger important objectives for the coming decade. Improving how we predict crop yields is one of the best methods to address these issues and raise food production rates. When we know what the crop harvest will be, we can adjust agriculture to achieve better food security.

Output from farm crops is essential when farmers must draught a plan for harvesting. The system looks at data about the soil, including its nutrients and what crops require to decide on the most appropriate choices for particular areas. More and more, machine learning is helping farmers process their data to give them crop suggestions that ensure higher production and greater profits. Upon using data analysis, farmers achieve higher productivity and face fewer risks because of crop failures.

Besides that, plants are central to both the natural environment and the economy. Even so, they are at risk from diseases that develop because of bacteria, fungi and viruses. Identifying these diseases fast is necessary to stop them from doing major harm. In this paper, a plant disease detector is presented, one that can detect plant diseases from images of leaves. With the help of CNN, the model is able to recognise several plant diseases accurately. The technology can recognise and alert farmers to diseases on the spot, so they can address the problem without delay.

1.1 Problem Statement and Proposed Solution

Farmers are dealing with different problems when selecting crops, handling fertilisers and stopping diseases. To deal with these issues, we are presenting an agricultural system that links together crop recommendations, suggestions for fertiliser use and plant disease checking. Because of machine learning and neural networks, the system is designed to support farmers in making decisions that reduce waste, help the farm work more efficiently and encourage sustainable practises in agriculture.

The system would help farmers manage their crops with precise advice on which crop to grow, how much fertiliser to add and what diseases might threaten them. Leveraging both machine learning and deep learning methods such as Random Forests, Decision Trees and Gaussian Naive Bayes, along with CNNs, will help us make choices supported by reliable information. The system will review soil nutrients, weather records and plant signs to give immediate suggestions to help farmers improve how they grow crops.

1.2 Scope of Project

The effort is centred on designing, installing and assessing an overall agricultural decision support system. The project has been organised into three distinct modules.

The Crop Recommendation Module studies soil nutrients (N, P, K) and other factors in the environment such as temperature, humidity, pH and the amount of rainfall, to recommend the most suitable plants for different regions. This work also means discovering and testing algorithms that work well across many types of agricultural settings and crop varieties. In addition, the system will use data gathered in real time to change crop recommendations if there are environmental changes.

Fertiliser application will be improved by the Suggestion Module that reviews soil nutrient data and crop needs. Programmes powered by machine learning will develop special plans for fertilising fields, boosting soil health while increasing how much we can grow and do so without harming the environment. Tests will be carried out in local fields to measure if the new fertiliser plans increase soil fertility and help reduce nutrients flushed out with water.

This module will predict plant diseases early on by analysing images of plant leaves with CNNs. The model will work with datasets, including Plant Village and its images cover both healthy and sick plants. Using sensors or smartphones to capture images, the system will help farmers get quick diagnoses and decide fast on measures to protect crops. Ensuring accuracy, sensitivity and specificity are key points of investigation for our CNNs.

All in all, this project hopes to plan, run and cheque an agricultural system that strengthens farmer productivity, cuts down on wasted resources and promotes sustainable growing. The system will work well for groups both large and small, be easy to use and always deliver useful insights into managing crops, fertiliser use and disease detection. The project makes use of data analytics, machine learning and deep learning to help overcome key concerns in farming and improve food supply everywhere.

2. Preliminary Work

This part of the document defines essential concepts that form the basis of the work, so readers are prepared for the models discussed later.

2.1 The description of time series data follows.

Time series data contains information taken at equally spaced intervals through time, just like stock prices or temperatures. Trend recognition, future prediction and finding uncommon patterns are crucial activities. The field of finance and also weather forecasting, depends heavily on time series analysis to guide choices. Data is first filtered and then studied with proper methods to achieve more accurate results.

A number of systems depend on time series data such as:

1. **Crop Suggestion System**
   a. At least once a month or once per season
   b. Properties measured: Soil pH, Temperature, Humidity and Rainfall
   c. Pick: The best crops for the season ahead

2. **The Fertilizer Suggestion System**
   a. You can apply this for the season or after each shot.
   b. Factors: Nutrients in the soil, Species of the crop
   c. Target: What fertilizer should you use and how much should you put down

3. **A System for Detecting Plants Diseases**
   a. Time Interval: Keeping an eye on the network system with occasional inspection
   b. Features include images of plants.
   c. Thing diagnosed: Disease

All data in each system is given a time stamp. Thus, the crop recommendation system checks Rainfall, Temperature and soil nutrient levels, whereas the plant disease detection system saves images. With synthetic data or data drawn from earlier records, experts are able to tune models and improve what they forecast.

2.1.1 covers the different types of Time Series Analysis.

Many methods are used in time series analysis, with each designed for a specific reason. Significant types of these are:

- It involves breaking up data into different groups.
- A curve is fit to the data to expose the relation between the variables.
- In Descriptive Analysis, you look for recurring trends and cycles.
- Explanatory Analysis is aimed at finding out what makes things happen.
- **Exploratory Analysis:** Shows certain important aspects of the data, using pictures when possible.
- **Forecasting:** Uses information from the past to predict coming values.
- Intervention Analysis looks at the effects of events on the collection of data.
- Segmentation means organizing data into smaller groups to study the structure beneath.

2.2 Data Preprocessing

Because crop recommendation data and plant disease data are not the same, preprocessing them for the two systems needs special treatment. The next section explains how each system goes through preprocessing.

1. **Crop Recommendation System – Provided in Numerical Form**
   - It is necessary to detect and care for missing values in the dataset. Solving the problem of missing values may be achieved by using techniques such as imputing with the mean, median or even KNN.

   - By using min-max scaling or standardisation, all numerical features (nitrogen (N), phosphorus (P), potassium (K), temperature, humidity, pH and rainfall) can be brought to the same scale. As a result, each property is represented fairly, since the level of detail varies little from one feature to another.

- Sometimes, it's necessary to make new features out of old ones to help the model make better decisions. A feature for soil fertility might be made by counting the levels of N, P and K together.

- If you have a categorical feature (such as soil type), convert it to a number using either one-hot encoding or label encoding which improves the quality of your model input.

- It is necessary to choose the features that play a big role in crop recommendation. By doing correlation analysis, importance ranking or by applying what's known in the field, you can find the most important features.

- After that, the complete dataset is broken up into two parts: training data and testing data. The purpose of this division is to prepare models and cheque how they handle data they have not seen yet.

2. **I examine the plant images within the Plant Disease Detection dataset.**
    - All images in the dataset are made the same size during the Image Preprocessing stage. Ensuring model efficiency means pixel normalisation, so values are adjusted from 0 up to 1. In addition, inducing common types of data transformation helps the model to recognise more kinds of features in new data, allowing it to perform better.

    - The dataset is expanded by applying activities such as rotating, translating, shearing and zooming to its images randomly. Because of this, the training dataset becomes more inhomogeneous which sharpens the model's robustness.

    - Label Encoding: To use the data, each disease must be assigned a label and that label must be encoded. These disease categories can be marked as one-hot or label encodings when working with numbers.

    - Right before training, the image dataset should be chopped into training, validation and test sets. The number of images in each set should represent the same number of classes (e.g., equal numbers of healthy and diseased plants).

    - Stacking – Transfering large image collections is essential by using Data Loader. Batching images during every stage of model development can be made easier by using loaders or generators.

    - Performing these preprocessing processes ensures that both crop recommendation and plant disease detection systems will be set up properly for model building and training.

2.3 Grouping Data

The act of data segmentation breaks a dataset into smaller pieces using chosen classifiers. For both the crop recommendation and plant disease detection systems, how segmentation works changes because of the difference in their data.

1. **In the Crop Recommendation System, I used numerical data.**
    - By organising the dataset according to time (year or season), one can record any seasonal Varies in agriculture factors. Time-based segmentation allows us to see patterns that last a long time or repeat every so often.

    - If we divide the data by location, it allows us to see how different soil, climate and agricultural practises impact things. As a result of this segmentation, guidelines can be fit to the special needs in every region.

- When we segment by crop type, recommendations can be customised for each crop. As a result, this process lets the system suggest farm tips and practises that are just right for the crop in question.

- Grouping data by soil pH, temperature or rainfall is how feature-based segmentation helps find patterns occurring within those groups of data.

- Interesting results about soil, climate or growth can be found by stratifying your samples by levels of fertility, various climates or different stages of crop growth.

2. System for Detecting Plant Disease with Imagery (Data Input).

- Image information is used in the plant disease detection system, where segmentation is shaped by factors including different disease types and their severity.

- When the data is separated by disease types (such as fungal, bacterial or viral diseases), it becomes possible to build accurate detection models for each.

- Dividing the data by severity of disease (mild, moderate or severe) allows disease strategies to match the level of infection and set priorities.

- Grouping data by crop type or plant helps recognise which crops may get diseases and how their symptoms change, permitting systems for detecting each type of crop.

- By splitting images by quality, you'll be able to stop low-quality images that tend to negatively influence model results.

- Once a dataset is split by sections such as leaves, stems or fruits, specialists can make models that look for diseases in each part of the plant.

- Such segmentation lets researchers collect consistent and reliable groups that improve the reliability of their models for both plant guidance and disease recognition systems.

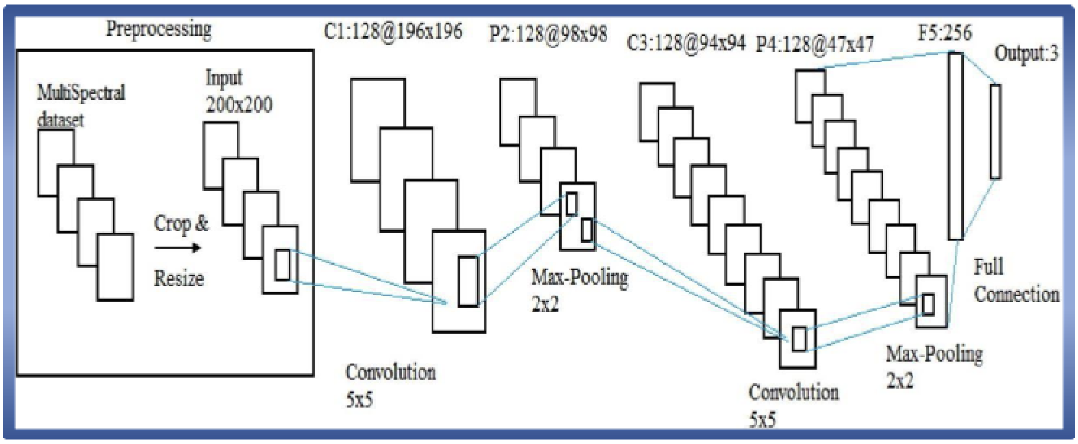

Fig1: CNN-Model

3. Requirement Analysis

Together, soil samples, information on crops and climate analysis, along with machine learning, can recommend fertiliser and notice diseases for better crop results. Algorithms are used in the system to choose fertilisers to match the soil, detect plant diseases fast and advise on the best management methods.

- Together, soil studies, information about crops and knowledge of climate are used by machine learning in the Integrated Agricultural System to improve how crops are cared for and their harvests. It includes:

- The Crop Recommendation System suggests suitable crops by studying details in soil, the climate and harvest history and uses Random Forest and Decision Trees to help.

- Fertiliser Suggestion System: Matches fertilisers to the needs of the crops as revealed by the soil, type and environment.

- With this system, CNN technology is used to recognise plant diseases in photos and offers suggestions for care.

This system delivers rapid processing, the ability to support a growing number of users, strong safety measures and healthy user interfaces. It works to achieve compliance by designing the system wisely, ensuring it is used by employees and making sure pieces of technology are compatible.

4. Implementation

All the necessary scripts, configuration files and documentation can be found with the GitHub project code and files. If you'd like to be involved, you are welcome to copy the repository andpersonalize it as you see fit.

4.1 Software Requirements

Requirement	Description
Python	Primary programming language, with the latest stable version being Python 3.11.0, featuring performance improvements and enhanced error reporting.
Jupyter Notebook	A local environment for faster execution without uploading files repeatedly.
Google Colab	An online environment that does not require installation and can be accessed via a Google account.

TABLE1: SOFTWARE-REQ

4.2 Libraries and APIs Used

Library/API	Description
NumPy	A library for numerical computing in Python, supporting multi-dimensional arrays and matrices.
Pandas	A data manipulation and analysis library, particularly for handling time-series and tabular data.
Matplotlib	A visualization library for creating various types of plots and charts.
Scikit-learn	A machine learning library providing algorithms for classification, regression, and clustering.
Keras	A high-level neural networks API, built on top of TensorFlow, for fast experimentation with deep learning models.
TensorFlow	A machine learning library focused on deep learning, developed by Google.
OpenCV	A library for real-time computer vision tasks, supporting GPU acceleration for efficient processing.

TABLE2: LIBRARIES & API

4.3 Validation Accuracy

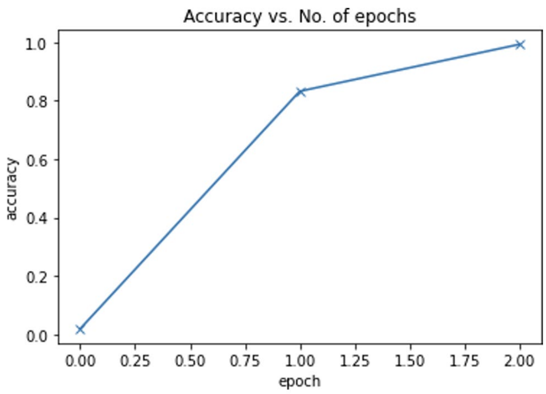

Fig2: validation-Accuracy

4.4 Validation Loss

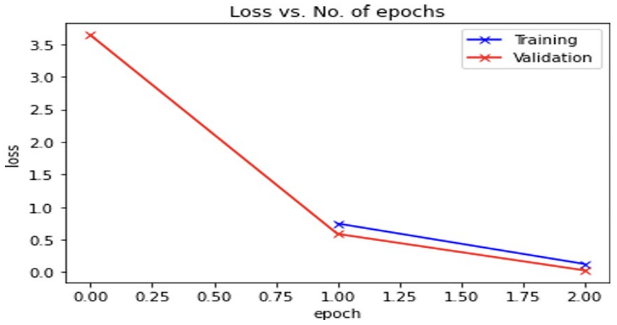

Fig3: validation-Loss

4.5 Learning Rate overtime

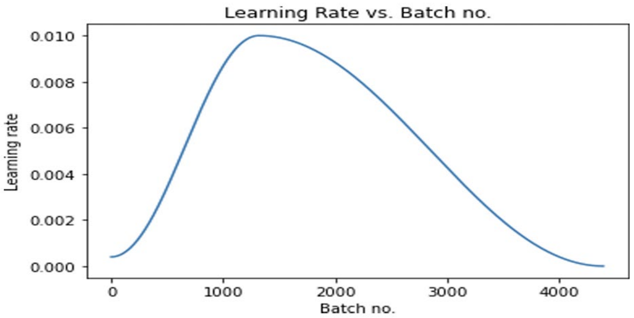

Fig4: Learning Rate

5. Conclusion

The system's key subsystems—Crop Recommendation, Fertiliser Suggestion and Plant Disease Detection—have brought big changes to agriculture, raising crop yield and quality. Random Forest, Decision Tree and Gaussian Naive Bayes algorithms are used in the Crop Recommendation System to give people personalised crop advice that considers both environmental data and records from history. With carefully considered soil, climate and geographical analysis, this system helps farmers make choices that raise their output and the profits from their work.

To go with this, the Fertiliser Suggestion System matches fertilisers to what crops and soil require. Getting soil nutrient data about the crop helps you use just enough fertiliser which benefits the environment and cuts down on waste.

Moreover, the system applies deep learning techniques, using a CNN ResNet model, to spot plant diseases from the pictures it examines. When diseases are found early, farmers can take steps to stop the spread, ensure healthy crops and raise both their output and the safety of the farm.
With these three parts, farmers have the guidance to boost their performance, watch their resource use and address problems from pests, diseases and climate impacts on their farms. Using modern methods and technology, the system shows important progress towards making farming more environmentally and economically friendly. Apart from helping the industry develop, it also supports a dependable and environmentally friendly source of food for later generations.

Acknowledgement

I am thankful to the Management of Amrita Sai Institute of Science and Technology for giving me an opportunity to work with his project.
I would like to the thank **Dr. M. Sasidhar**, Principal, Amrita Sai institute of science and technology, for his constant encouragement and support during the progress of this work.
I am deeply grateful to **Dr. P. Chiranjeevi**, Professor and Head of the Department, for his valuable guidance and consistent support during the course of the project.
A special note of thanks to my internal guide, **Mrs. Y. Nirmala (M.Tech)**, for her exceptional guidance, constant motivation, and continuous encouragement, which played a crucial role in the successful completion of this project.

NAYENI VAMSI

References

[1] C. Chen and J. Fu, "Construction & application of knowledge decision tree after a disaster for water body information extraction from remote sensing images," *Journal of Remote Sensing*, vol. 2, no. 2, pp. 792-801, 2018.

[2] W. Lui and Z. Wang, "A survey of deep neural networks architecture," *Neurocomputing*, vol. 23, no. 2, pp. 25-37, 2019.

[3] B. Mamandipoor and M. Maid, "Monitoring & detecting faults in wastewater treatment plants using deep learning," *Environmental Monitoring and Assessment*, vol. 192, no. 2, p. 148, 2020.

[4] G. Zhao, Z. Xu, and B. Pang, "An enhanced inundation method for urban flood hazard mapping at the large catchment scale," *Journal of Hydrology*, vol. 571, no. 2, pp. 873-882, 2019.

[5] S. Minaee and Y. Wang, "An ADMM approach to masked signal decomposition using subspace representation," *IEEE Transactions on Image Preprocessing*, vol. 28, no. 28, pp. 3192-3204, 2019.

[6] Y. Lia and G. Juan, "Flood monitoring system using DCNN," *Neurocomputing*, vol. 2, no. 4, pp. 236-248, 2019.

[7] A. Davrancvhe, G. Lefebvre, and B. Poulin, "Wetland monitoring using classification trees and SPOT-5 seasonal time series," *Remote Sensing of Environment*, vol. 114, no. 2, pp. 552-562, 2012.

[8] L. Y, H. Qin, and Z. Zhang, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, no. 2, pp. 11-26, 2017.

[9] T. A and L. Dong, "Identification of water bodies in Landsat8," *Sensors*, vol. 1075, no. 2, p. 16, 2016.

[10] J. Marcais, D. Dreuzy, and J. R, "Prospective interest of deep learning for hydrological inference," *Groundwater*, vol. 55, no. 5, pp. 688-692, 2017.