Application of Intelligent Recommendation for Agricultural cultivation.

**Abstract**

Farmers face several challenges when growing crops like uncertain irrigation, poor soil quality, etc. Especially in India,a major fraction of farmers do not have the knowledge to select appropriate crops and fertilizers. Moreover, crop failure due to disease causes a significant loss to the farmers, as well as the consumers. While there have been recent developments in the automated detection of these diseases using Machine Learning techniques, the utilization of Deep Learning has not been fully explored. Additionally, such models are not easy to use because of the high-quality data used in their training, lack of computational power, and poor generaliz ability of the models. To this end, we create an open-source easy-to-use web application to address some of these issues which may help improve crop production.In particular, we support crop recommendation, fertilizer recommendation and plant disease prediction .In addition, we also use inter pretability techniques in an at tempt to explain the prediction made by our disease detection model

**Chapter 1**

INTRODUCTION

Agriculture, as we all know, is the foundation of the Indian economy. Agriculture is an important occupation in India. More than 60% of the country's land is used for agriculture, which feeds 1.3 billion people . Agriculture is the cultivation of plants and animals. In India, agriculture gave rise to civilization. We need soil to cultivate crops. As a result, soil is a critical factor in agriculture. Soil health is essential for good food production. It provides the roots with essential nutrients, water, oxygen, and support. Soil is the foundation of the food system, as well as the location of all plants used in food production. In India, several soil varieties are available. They are alluvial soil (cotton, rice), black soil (sugarcane, sunflower), red soil (corn, ragi), laterite soil (pulses, tea, coffee), and so on. Many studies have been conducted to improve agricultural planning. The crop can be recommended using a machine learning technique.

Machine learning is an subfield of artificial intelligence that describes a machine's ability to mimic intelligent human behavior. Artificial intelligence systems are employed in the same way as humans do to automate complex tasks . Machine learning begins with data, such as financial transactions, individuals, or photos.

The information is collected and processed to be utilized as training data for the machine learning system. If the data is more then the software shows better results. After that, the developer select a ML model to use, input the data, and train the system to find patterns or make predictions on its own.

The environmental data that is gathered by remote sensors are processed by algorithms and statistical data which will be understood and helpful to farmers for decision makings and keep track of their farms. The more inputs and statistical data collected, and higher the algorithmic rule is at predicting the outcomes. And the aim is that farmers will use these technologies to attain their goal of improved harvest by creating better selections within the field. By implementing the system of temperature, soil hydrogen ion concentration and soil wetness detection, the information captured are processed with an explicit algorithmic rule and passed to a centralized database that is connected to different modules of the research, so the main system will predict the most effective crop kind that the farmer should grow to require the most outcome of the crop kind that is farmed in a home garden or the respectable land area

objective:

* Farming is one of the major sectors that influences a country’s economic growth.
* In country like India, majority of the population is dependent on agriculture for their livelihood. Many new technologies, such as Machine Learning and Deep Learning, are being implemented into agriculture so that it is easier for farmers to grow and maximize their yield.

**Chapter 2**

LITERATURE REVIEW

A study of machine learning algorithms was conducted in a research paper by Rashi Agarwal . This system would help farmers make educated decisions about which crops to grow based on a variety of environmental and geographical factors. They employed decision trees, KNNs, Random Forests, and neural networks. The neural network had the highest accuracy of all of them.

Priyadharshini A conducted a study on machine learning algorithms in her research article. This technology reduces crop failure and decreases productivity by supporting farmers in choosing the proper crop and provide the data that regular farmers do not maintain. A variety of machine learning algorithms were applied. The neural network was the most accurate of the bunch.

Shilpa Mangesh Pande In her research article, she presents a farmer-friendly and realistic production forecasting system The suggested technology is connected to farmers via a mobile application. The user's location is determined with the help of GPS. All of the algorithms are compared in terms of crop yield forecast accuracy. The RF algorithm showed to bethe best for the provided data set, with a 95% accuracy.

Mayank Champaneri conducted research on crop yield prediction using a data mining technique. They used a random forest classifier because it can perform classification and regression tasks. The user-friendly website built that can be used by anyone to predict crop yield for their choice of crop by giving climate data for that area.

In this literature review of this project, the team sought out and studied various patents, research papers, documents, and newspapers and magazine articles from various scenes. The paper states requirements and why they tend to move into precision agriculture which is due to globalization are discussed. Precision agriculture is site-specific farming. Though Precision agriculture has shown an improvement with time, there exist some issues. As mentioned above site-specific methods of such systems are needed to be supervised to get an improved result. Only a few of the outcomes are provided a particular result. Nevertheless, the situation is farming is indispensable since if any default or a mistake occurs, it might lead to serious damage to resources and as well as the plants.

In this research, it is proposing a system where the major factors are taken into consideration at the same time and come up with a solution so that the system will not be complicated for the user. As mentioned above in the sentence, the major factors taken into consideration at once is unlike other models proposed in previous researchers, this system considers all the major factors which are essential for plant growth, are processed together using various algorithms whereas the other models consider only parameters at once keeping the other factors constant

Chapter 3

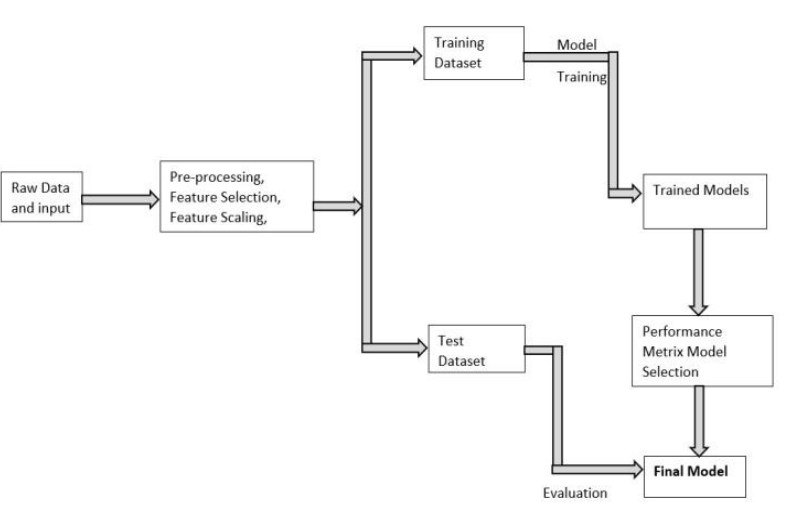
PROPOSED SYSTEM

* In our project, I present a website in which the following applications are implemented; Crop recommendation, Fertilizer recommendation and Plant disease prediction, respectively.
* In the crop recommendation application, the user can provide the soil data from their side and the application will predict which crop should the user grow.
* For the fertilizer recommendation application, the user can input the soil data and the type of crop they are growing, and the application will predict what the soil lacks or has excess of and will recommend improvements.
* For the last application, that is the plant disease prediction application, the user can input an image of a diseased plant leaf, and the application will predict what disease it is and will also give a little background about the disease and suggestions to cure it.

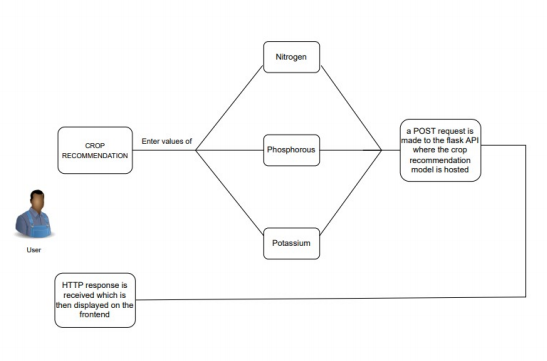
Advantage:

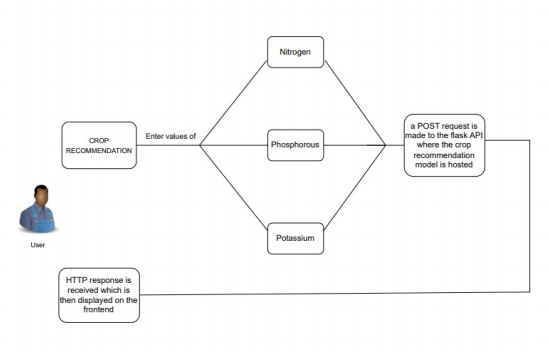
* ML and DL systems are helping to improve the overall harvest quality and accuracy – known as precision agriculture.
* this technology helps in detecting disease in plants, pests and poor nutrition of farms.
* sensors can detect and target weeds and then decide which herbicide to apply within the region.

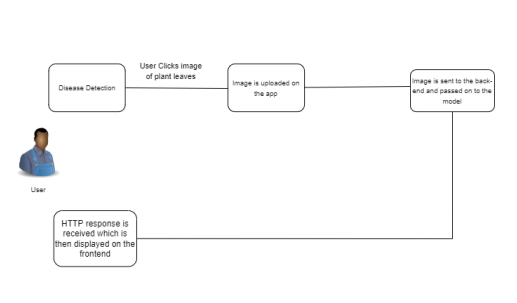
Block Diagram:



**FLOW DIAGRAM FOR CROP**







METHODOLAGY

* Crop Recommendation

In the authors use - Random forests, Artificial NeuralNets, Support Vector Machines, etc. and conclude that Random forests work best for their dataset in crop recommendation.They also create a mobile application system which takes in location data using GPS and predicts the crop yield for a givencrop, in addition to recommending crops based on area and oil quality as input. Random Trees for crop recommendation.

* Fertilizer Recommendation

A lot of research has been done in fertilizer recommendation and a majority of them use the N, P, K, pH values ofsoil sometimes in addition with depth, temperature, weather,location, precipitation. The usual approach is to use rule based classification, but some approaches also use clusteringon fertilizer data using K-Means and Random Forests for recommendation.

Plant Disease Detection has been a very active field of research and there are several different techniques which have been proposed over the years, the latest ones using deep learning approaches.

Dataset

The dataset for this topic was taken from Kaggle. It is acrop recommendation dataset giving us information abou tvarious types of crops and the features that decide which crop is suitable for growing

Features of the Dataset

N: ratio of Nitrogen content in soil

P: ratio of Phosphorous content in soil

K: ratio of Potassium content in soil

Temperature: temperature in degree Celsius

Humidity: relative humidity in %

pH: pH value of the soil

rainfall: rainfall in mm

MODULES:

* DataSet
* Data preprocessing.
* Splitting dataset.
* Build the model
* Training the model.
* Evaluating the model.
* Testing the model
* Saving the model.

Dataset collection

* Collecting data allows you to capture a record of past events so that we can use data analysis to find recurring patterns.
* From those patterns, you build predictive models using machine learning algorithms that look for trends and predict future changes.
* Predictive models are only as good as the data from which they are built, so good data collection practices are crucial to developing high-performing models.
* The data need to be error-free (garbage in, garbage out) and contain relevant information for the task at hand.
* For example, a loan default model would not benefit from tiger population sizes but could benefit from gas prices over time
* In this module, we collect the crop recommendation data from kaggle dataset archives. This dataset contains the information of divorce in previous years.

Data preprocessing.

Data Cleaning

* Data cleaning is a critically important step in any machine learning project.
* In this module data cleaning is done to prepare the data for analysis by removing or modifying the data that may be incorrect, incomplete, duplicated or improperly formatted.
* In tabular data, there are many different statistical analysis and data visualization techniques you can use to explore your data in order to identify data cleaning operations you may want to perform.

Feature Extraction:

* This is done to reduce the number of attributes in the dataset hence providing advantages like speeding up the training and accuracy improvements.
* In machine learning, pattern recognition, and image processing, feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations. Feature extraction is related to dimensionality reduction
* When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a reduced set of features (also named a feature vector).
* Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.

Model training

* A training model is a dataset that is used to train an ML algorithm. It consists of the sample output data and the corresponding sets of input data that have an influence on the output.
* The training model is used to run the input data through the algorithm to correlate the processed output against the sample output. The result from this correlation is used to modify the model.
* This iterative process is called “model fitting”. The accuracy of the training dataset or the validation dataset is critical for the precision of the model.
* Model training in machine language is the process of feeding an ML algorithm with data to help identify and learn good values for all attributes involved.
* There are several types of machine learning models, of which the most common ones are supervised and unsupervised learning.
* In this module we use supervised classification algorithms like linear regression to train the model on the cleaned dataset after dimensionality reduction.

Testing model:

* In this module we test the trained machine learning model using the test dataset
* In machine learning, model testing is referred to as the process where the performance of a fully trained model is evaluated on a testing set.
* in machine learning, a programmer usually inputs the data and the desired behavior, and the logic is elaborated by the machine.
* therefore, the purpose of machine learning testing is, first of all, to ensure that this learned logic will remain consistent, no matter how many times we call the program in the testing

Performance Evaluation

* In this module, we evaluate the performance of trained machine learning model using performance evaluation criteria such as F1 score, accuracy and classification error.
* In case the model performs poorly, we optimize the machine learning algorithms to improve the performance.
* Performance evaluation is an important aspect of the machine learning process. However, it is a complex task.
* It, therefore, needs to be conducted carefully in order for the application of machine learning to radiation oncology or other domains to be reliable.
* used to evaluate a classification model are accuracy, precision, and recall is used crop recommendation

Prediction

* Prediction” refers to the output of an algorithm after it has been trained on a historical dataset and applied to new data when forecasting the likelihood of a particular outcome, such as whether or not a customer will churn in 30 days.
* The algorithm will generate probable values for an unknown variable for each record in the new data, allowing the model builder to identify what that value will most likely be.
* The word “prediction” can be misleading. In some cases, it really does mean that you are predicting a future outcome, such as when you’re using machine learning to determine the next best action in a marketing campaign.
* Other times, though, the “prediction” has to do with, for example, whether or not a transaction that already occurred was fraudulent.
* In that case, the transaction already happened, but you’re making an educated guess about whether or not it was legitimate, allowing you to take the appropriate action.
* In this module we use trained and optimized machine learning model to predict whether the crop recommendation

**UML DIAGRAM**

Unified Modelling Language (UML) is simply another graphical

representation of a common semantic model. The proposed system has been

designed by using use case diagram, class diagram, sequence diagram,

collaboration diagram, state chart diagram and component diagram.

**USE CASE DIAGRAM**

The use case diagram consists of the actors and the use cases. The actors

of the system are user, system holder, device controller and the use cases are authentication, checking credentials, basic ON/OFF, allow/deny user, storing NLP commands, Input through voice commands, Deriving Data, Intrusion Detection, Service Maintenance. Figure No. 3.2 describes Use Case diagram for Adaptive Automation System (AAS).

User and Non-User can login to the device by Admin. Then device controller decides the entry of the user. The user can control the devices through voice commands and these can be monitored by Device Controller. Intrusion detection and maintenance also controlled by the Device Controller (Admin). The user can check for the intruder in the environment and based on the intruder the door can be open/closed. The light can be turned on/off by the system holder and they can alert the emergency system if any poisonous gas or temperature exceeds the normal level. This can be maintained by system maintenance.

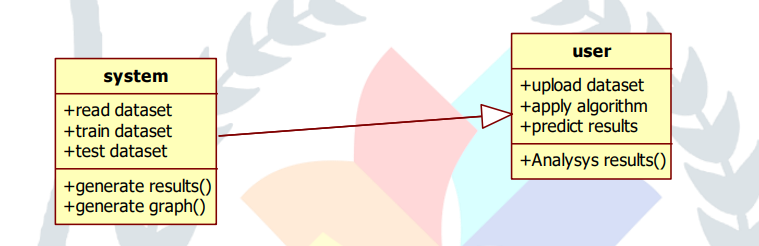


**CLASS DIAGRAM**

Class diagram is to model the static view of an application. Class diagrams are the only diagrams which can be directly mapped with object-oriented languages and thus widely used at the time of construction and it is used for general conceptual modelling of the structure of the application, and for detailed modelling translating the models into programming code.

Class diagrams can also be used for data modelling. The classes in a class

diagram represent both the main elements, interactions in the application, and the classes to be programmed.



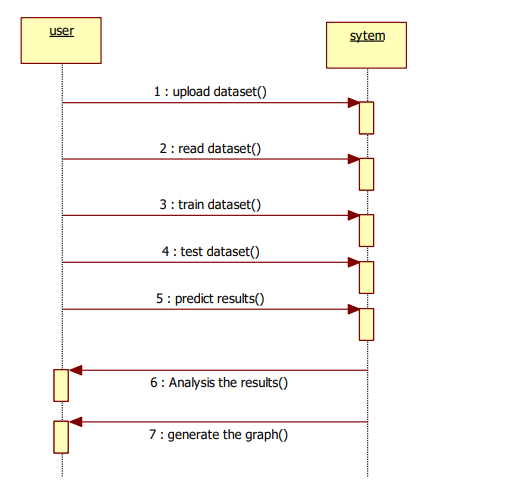
**SEQUENCE DIAGRAM**

The control flow between various participants or entity roles of the

corresponding system in the form of messages is represented in the Sequence Diagram. The participants are represented within the rectangular object. The swim line or the lifeline that is dragged below every participant represents the lifetime of the corresponding participant.

The UML representation of a class is rectangle containing three

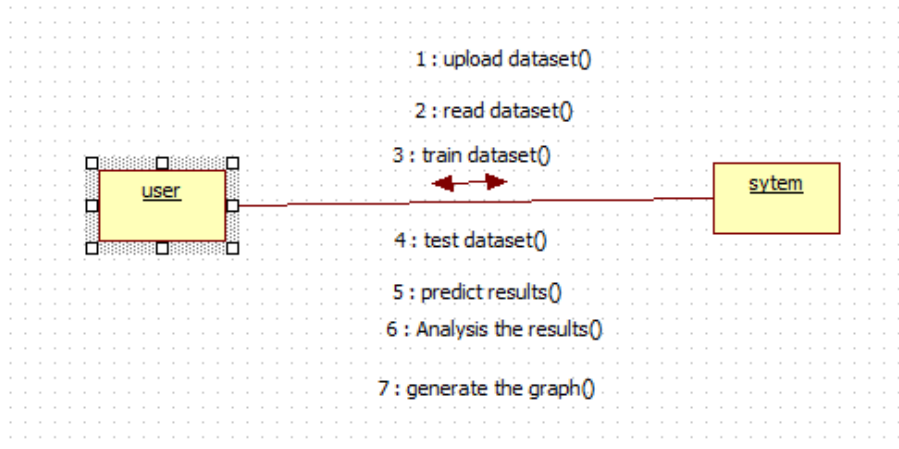
compartments stacked vertically. The top compartments shows the class’s name. the middle compartments list the class’s attributes. The bottom compartment lists the class operations known as the methods of the class. A class diagram consists of any number of classes which will be connected by the lines, which may have arrows at one or both ends, connecting the boxes. These lines define the relationships, also called associations, between the classes. These lines will have multiplicity to represent the number of instances of the classes.



**COLLABORATION DIAGRAM**

Collaboration diagram is defined as one of the interaction diagram,

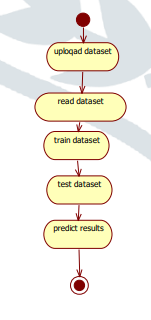
which consists of the set of objects related in a particular context and interaction among those objects. The collaboration diagram is also called as the set of message exchange among the objects within the collaborative nature of message exchange between the corresponding objects.



**ACTIVITY DIAGRAM**

Activity diagram is another important diagram in UML to describe the

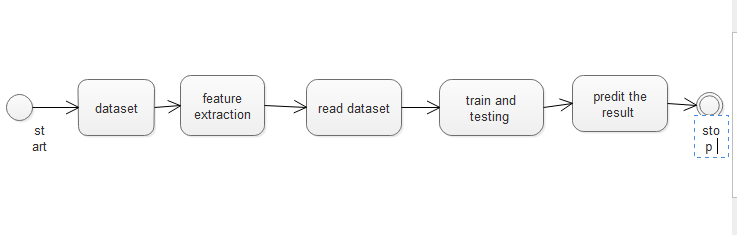
dynamic aspects of the system. Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The control flow is drawn from one operation to another with different components of activity diagram. Some of the components of activity diagram Start/Stop symbol, Action symbol, Joint and Fork symbol, Decision symbol, Connector symbol.



**STATE CHART DIAGRAM**

State chart diagram is one of the five UML diagrams used to model the

dynamic nature of a system. They define different states of an object during its lifetime and these states are changed by events. State chart diagrams are useful to model the reactive systems. There are three main components in the State chart diagram such as Initial/ Final states, State symbol and Transition symbol



**COMPONENT DIAGRAM**

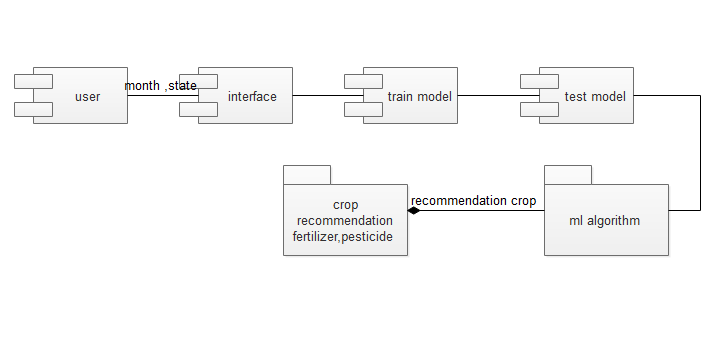
Component diagrams are used to visualize the physical components in a

system. These components are libraries, packages, files, etc. Component

diagrams can also be described as a static implementation view of a system. The basic notations of the component diagrams are such as Component and Connector.

These diagrams show the physical components of a system. To clarify it,

we can say that component diagrams describe the organization of the components in a system. Organization can be further described as the location of the components in a system. These components are organized in a special way to meet the system requirements.



**TECHNICAL BACKGROUND**

**LIST OF MODULES**

The overall proposed system is divided into three modules. This division is made based upon the development strategies and the time constraints that are in need for the project development. The first module which forms the base for the proposed system. The second module is used for achieving the initial commandbased master-slave working. The final module is used for achieving voice-based recognition and control of the sensors based upon the dynamic environmental conditions. The modules are as follows

Chapter 4

CONCLUSION

we propose a user-friendly web applicationsystem based on machine learning and web-scraping calledthe ‘Farmer’s Assistant’. With our system, we are successfullyable to provide several features - crop recommendation using Random Forest algorithm, fertilizer recommendation using a rule based classification system, and crop disease detection using EfficientNet model on leaf images. The user can provide the input using forms on our user interface and quickly get their results. In addition, we also use the LIME interpretability method to explain our predictions on the disease detection image, which can potentially help understand why our model predicts what it predicts, and improve the datasets and models using this information

REFERENCES

* S. D. Khirade, A. B. Patil, “Plant Disease Detection Using ImageProcessing”, 2015 International Conference on Computing Communi\_x0002\_cation Control and Automation, 2015, pp. 768-771, doi: 10.1109/IC\_x0002\_CUBEA.2015.153.
* A. V. Panchal, S. C. Patel, K. Bagyalakshmi, P. Kumar, I. RazaKhan, M. Soni, “Image-based Plant Diseases Detection using Deep Learning”, Materials Today: Proceedings, 2021, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.07.281.
* J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database”, 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.
* Russakovsky et al., “ImageNet Large Scale Visual Recognition Chal\_x0002\_lenge”, International Journal of Computer Vision, 2015, 115(3), 211- 252. https://doi.org/10.1007/s11263-015-0816-y.
* George A. Miller, “WordNet: A Lexical Database for English”, Com\_x0002\_munications of the ACM, 1995, Vol. 38, No. 11: 39-41.
* A. Krizhevsky, I. Sutskever, G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks”, Part of Advances in Neural Information Processing Systems 25 (NIPS 2012)