VISVESVARAYA TECHNOLOGICAL UNIVERSITY
BELAGAVI

Project Report on

"NUTRIENT DEFICIENCY DETECTION IN PLANTS FOR FERTILIZER MANAGEMENT"

Submitted in the partial fulfillment for the requirements of the degree of

BACHELOR OF ENGINEERING IN
COMPUTER SCIENCE AND ENGINEERING

Submitted By

M AISHWARYA (1BY17CS085)

M MERLYN MERCYLONA (1BY17CS088)

NAMRATA KARKI (1BY17CS205)

Under the guidance of

Mr. Jagadish P
Assistant Professor
Department of CSE, BMSIT&M

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

BMS INSTITUTE OF TECHNOLOGY & MANAGEMENT
YELAHANKA, BENGALURU - 560064.

2020-2021

VISVESVARAYA TECHNOLOGICAL UNIVERSITY BELAGAVI

BMS INSTITUTE OF TECHNOLOGY &MANAGEMENT
YELAHANKA, BENGALURU – 560064

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the Project work entitled **"NUTRIENT DEFICIENCY DETECTION IN PLANTS FOR FERTILIZER MANAGEMENT"** is a bonafide work carried out by **M Aishwarya (1BY17CS085), M Merlyn Mercylona (1BY17CS088), Namrata Karki (1BY17CS205)** in partial fulfillment for the award of **Bachelor of Engineering Degree in Computer Science and Engineering** of the **Visvesvaraya Technological University, Belagavi** during the year 2020-2021. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in this report. The project report has been approved as it satisfies the academic requirements in respect of project work for B.E Degree.

Signature of the Guide	**Signature of the HOD**	**Signature of Principal**
Mr Jagadish P Assistant Professor Dept. of CSE, BMSIT&M	Dr.Bhuvaneshwari C.M Professor & HOD, Dept. of CSE, BMSIT&M	Dr. Mohan Babu G N Principal, BMSIT&M

External VIVA-VOCE

Name of the Examiners   Signature with Date

1)

2)

ACKNOWLEDGEMENT

We are happy to present this project after completing it successfully. This project would not have been possible without the guidance, assistance and suggestions of many individuals. We would like to express our deep sense of gratitude and indebtedness to each and every one who has helped us make this project a success.

We heartily thank our Principal**, Dr. MOHAN BABU G N, BMS Institute of Technology & Management,** for his constant encouragement and inspiration in taking up this project.

We heartily thank our **Head of the Department, Dr. Bhuvaneshwari C.M , Department of Computer Science and Engineering, BMS Institute of Technology & Management,** for her constant encouragement and inspiration in taking up this project.

We gracefully thank our Project Guide, **Mr. Jagadish P, Assistant Professor, Department of Computer Science and Engineering,** for his guidance, support and advice.

Special thanks to all the staff members of Computer Science Department for their help and kind co-operation.

Lastly we thank our parents and friends for the support and encouragement given to us in completing this precious work successfully.

M Aishwarya (1BY17CS085)
M Merlyn Mercylona (1BY17CS088)
Namrata Karki (1BY17CS205)

ABSTRACT

We attempt to use computational tools for nutrition monitoring as a part of decision support & farm management tools, which is of immense use to the farmers who do not have access to expert advice. In order to classify the deficiencies, we need to associate the symptoms of the affected plant with the nutrient that is deficient and causing those effects. We have developed an application that uses the leaf images of crops like rice, wheat and maize to identify nutrient deficiencies like Potassium, Magnesium, Zinc, Iron, Manganese, Copper, Boron, and Sulphur and recommends appropriate type and amount of fertilizer. It also considers some additional details like leaf age, other symptoms for diagnosis, which is performed using neural network model.

Hence using this application through mobile or web farmers can make informed decision to buy the adequate quantity of the fertilizer to treat their plants appropriately.

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Plants need nutrients to properly grow, thrive and survive. Hence it is essential that they receive the right type and quantity of nutrients at the right time. A plant needs almost twelve nutrients for its efficient growth. To achieve their life cycle and physiological functions, plant need chemicals such as Nitrogen, Potassium, Phosphorus, Calcium, Magnesium and Sulphur called the macronutrients and Zinc, Copper, Boron, Chlorine, Manganese, Iron called the micronutrients. These kind of nutrient deficiencies are generally identified in laboratories and through research. This can be a time consuming and costly process for the farmer to perform regularly. The accurate determination of nutritional status can prevent losses and also serve as a basis for the rational use of nutritional supplements. As a result, waste of resources are avoided and environmental impacts are also reduced. Moreover, computational tools for nutrition monitoring can be made available as part of decision support and farm management tools, which can be particularly valuable for farmers that do not have access to expert advice.

1.1 Background

Almost 70% of India’s population depends on agriculture, with 78% of them are small holder farmers. These farmers take loans to buy the requirements like fertilizers, pesticides, etc. for cultivation of their crops. When they notice some issue in their crops like leaf yellowing, leaf dropping, stunted growth, lack of flowering, etc. they ask their local fertilizer/pesticide shop for advice, and follow that. This method does not focus on identifying the specific cause of the problem, but rather treats the symptoms in a general manner. This can and has led to over-fertilization of soil which leads the chemicals to runoff along with the water, polluting the water bodies, causing eutrophication, etc. The nutrient deficiency in plants can disrupt their growth and food formation and result in poor flowering and fruiting. It can also lead to unwanted developments in the plant, where there is a large accumulation of one kind of nutrient, which prevents/blocks the intake of other nutrients. All this leads to loss in crop productivity, quality and eventually the income of the farmer due to low quality and poor yield, even though the farmer has spent money in buying the required fertilizers, leading them to debt and suicide.

Today, the majority of the farmers depend on their local fertilizer and pesticides shop, when they face some issues in their crops like leaf yellowing, leaf dropping, lack of blooms, etc. They blindly follow the recommendation given in the shop, without probing deeper to find the root cause of the problem and the reason for the symptoms. Various methods and extensive work has been conducted to identify nutrient deficiencies in plants. The nutrient deficiencies are identified in the leaves of the crop plants with the help of their symptoms like reduction in leaf size, changes in the color of the leaves, distorted edges, necrosis and black spots. In order for recognizing the deficient nutrients the tester may need to uproot the entire plant and carry out the manual work required. However with the advancement of technology several scientists and researchers have come up with the techniques to unload the heaps of work required for the identification. These include using the techniques from Machine Learning, Artificial Intelligence, Deep Learning, and Image Processing, where the color and texture features of leaves and fruits are used to identify deficiencies.

A plant needs almost twelve nutrients for its efficient growth. To achieve their life cycle and physiological functions, plant need chemicals such as Nitrogen, Potassium, Phosphorus, Calcium, Magnesium and Sulphur called the macronutrients and Zinc, Copper, Boron, Chlorine, Manganese, Iron called the micronutrients. These kind of nutrient deficiencies are generally identified in laboratories and through research. This can be a time consuming and costly process for the farmer to perform regularly. The accurate determination of nutritional status can prevent losses and also serve as a basis for the rational use of nutritional supplements. As a result, waste of resources are avoided and environmental impacts are also reduced. Moreover, computational tools for nutrition monitoring can be made available as part of decision support and farm management tools, which can be particularly valuable for farmers that do not have access to expert advice.

1.2 Literature Survey

G. Xu, et.al , proposed a method to identify Nitrogen, Phosphorus and potassium deficiency in tomato plants based on colour value features extracted [1].

K. A. M. Han, et.al, used convolutional neural networks to identify nutrient deficiencies in black gram plants, through leaf images. The nutrients identified were Calcium, Iron, Magnesium, Nitrogen, Potassium and Phosphorous [2].

C. A. Dhawale, et.al, used various image processing techniques like thresholding, segmentation, etc. to identify nutrient deficiencies in citrus plants [3].

U. Watchareeruetai, et.al, used convolution neural networks to identify nutrient deficiencies in black gram. The nutrient deficiencies were identified for the following elements: Calcium, Iron Potassium, Magnesium and Nitrogen [4].

M. Latte, et.al., propose a method of using HSV features in images to classify paddy plants as healthy or not healthy, and further identify nitrogen, phosphorus and potassium deficiencies, using a rule-based approach[5].

S. Sivagami, et.al used image processing techniques to identify the nutrient deficiency in tomato plants and also provides the corresponding diseases caused due to that deficiency [6].

1.3 Motivation

After the green revolution, there has been a huge rise in consumption of fertilizers and pesticide, as the hybrid crops require more resources like water and nutrients, as it lacks the qualities of normal indigenous plants. This leads the farmers to take loans to buy the chemicals, and are faced with debt when prices go low or harvest is poor. Additionally the chemicals in fertilizers are carried along with the water, when applied in excess and end up polluting water bodies and causing eutrophication. Moreover, when there is excess of one kind of nutrient in the soil that is taken up by the plant it blocks other essential nutrients from being taken up by the plant, leading to a deficiency. These excess chemicals end up in our food, leading to various health issues. So there is a need to apply the right kind and quantity of fertilizers to plants, whenever required. To tackle this issue, we propose a method of identifying the nutrient deficiencies in plants from its leaf images, to allow the farmer to make a more informed decision. During the last decade, the combination of digital images with artificial intelligence and image processing techniques for tackling agricultural problems has proven to boost the field of digital farming. A considerable effort in the development of new methods for detection of nutritional problems in plants can help in eliminating the problems and ensures to take the necessary measures for safe and sound growth of plants.

1.4 Problem Statement

After the green revolution, there has been a huge rise in consumption of fertilizers and pesticides this lead the farmers to take loans to buy the chemicals, and are faced with debt when prices go low or harvest is poor. Additionally the chemicals in fertilizers are carried along with the water, when applied in excess and end up polluting water bodies and causing eutrophication. Moreover, when there is excess of one kind of nutrient in the soil that is taken up by the plant it blocks other essential nutrients from being taken up by the plant, leading to a deficiency. These excess chemicals end up in our food, leading to various health issues. So there is a need to apply the right kind and quantity of fertilizers to plants, whenever required. To tackle this issue, we propose a method of identifying the nutrient deficiencies in plants from its leaf images, to allow the farmer to make a more informed decision

1.5 Aim and Objective

We propose to build a machine learning model that is capable of identifying the nutrient deficiencies in plants through leaf images. The farmer will just have to take a photo of the affected leaf, and provide some details like, age of plant, etc. and get the nutrient analysis for that plant. This knowledge in the hands of farmers will help them make a more informed decision when purchasing and applying fertilizers. The model can be deployed in the form of a mobile application, for easy use.

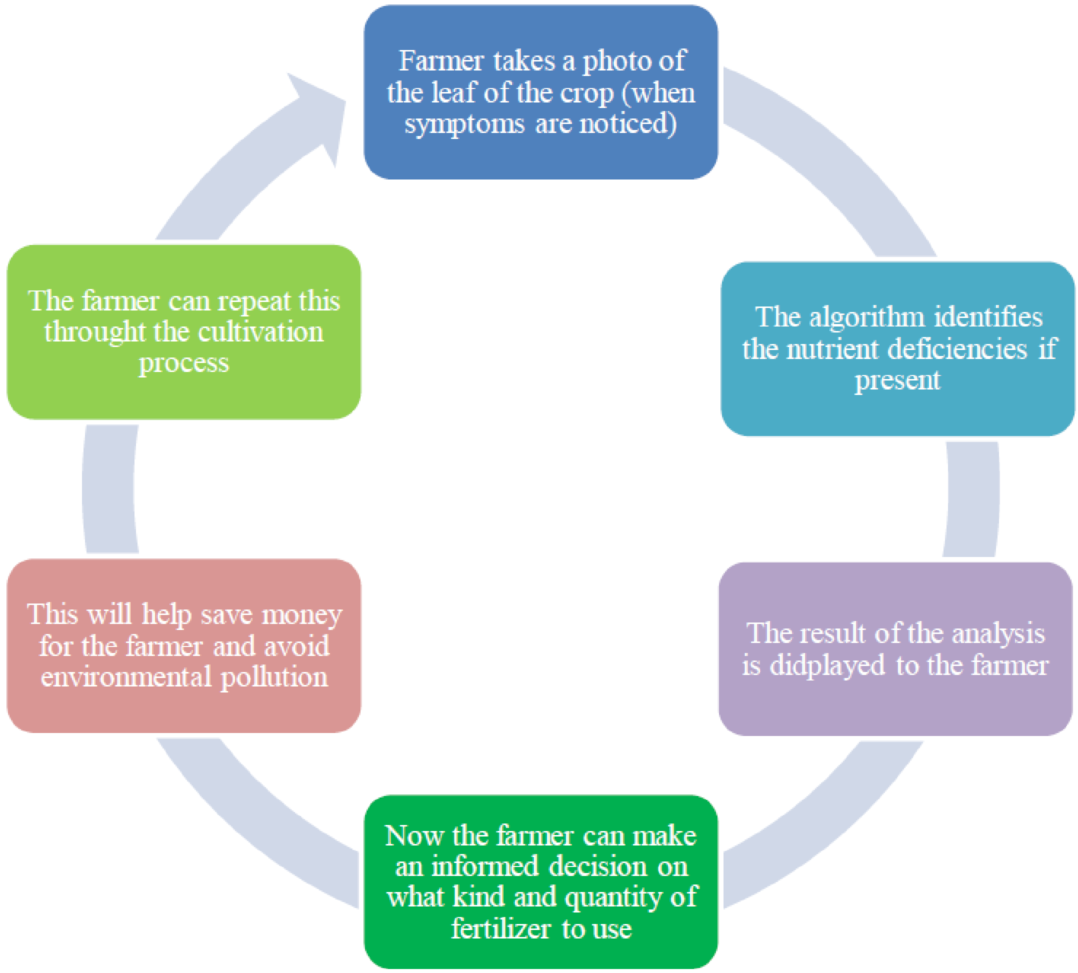

Fig 1.5.1 Aim and objective of the project

1.6 Scope of the Project

1) Reduce investment for farmers, thus saving money.
2) Ensure good productivity and yield in crops
3) Cheaper than performing multiple soil tests
4) Quick feedback in the field itself
5) Prevent over fertilization, and environmental pollution

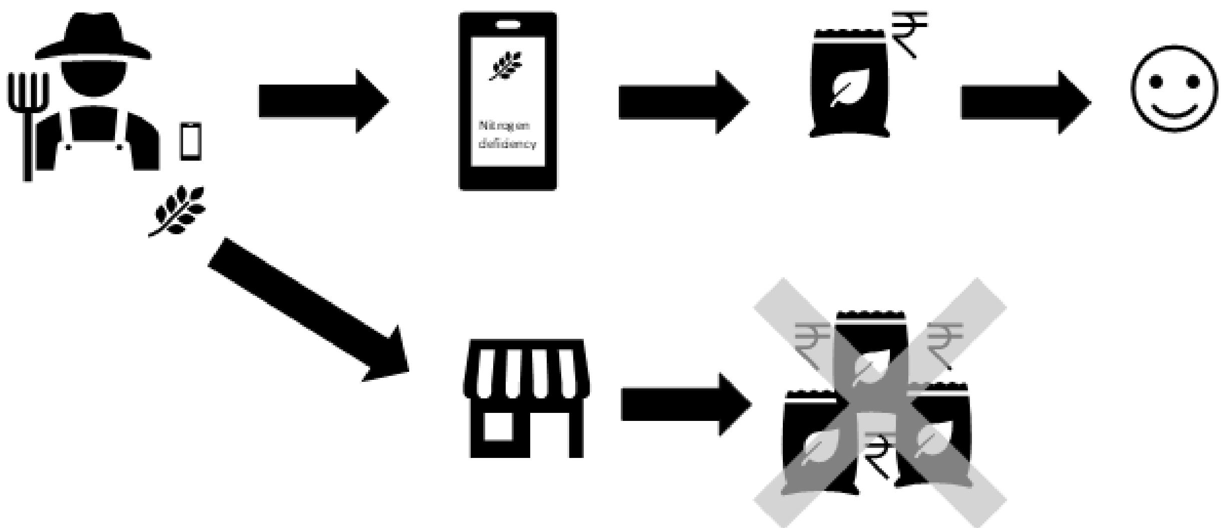

Fig 1.6.1 Scope of the project

1.7 Challenges

The existing methods largely focus on diagnosing macronutrient deficiencies i.e. Nitrogen, Calcium, Magnesium and Potassium etc. Minimal amount of work has been carried out for diagnosing micronutrient deficiency symptoms which also plays a role in causing serious damage to the growth of the plant. Hence, the future work should concentrate on diagnosing micronutrient deficiencies such as Iron, Zinc, Manganese, Boron, Chlorine etc. Also, other indicators of malady in plants should also be considered for future work and the data sets used must be created and widened to include other factors that might also be affecting the plant condition. Some of the problems include blight (death of plant part), bleaching (white coloration on leaves) or rust (formation of orange to reddish-brown spots) etc. Moreover the existing systems focus on a single crop, grown in a controlled environment, for a short period of time. Additional work is required to make use of technology in this field. Recently there has been an increase in usage of CNN and ANN for building classification models, this approach requires large number of datasets.

So there is a need for the proposed solution, to be scalable, and relevant to the actual symptoms encountered in the field.

1.8 Organization of the thesis

This project report is organised as follows:

- Chapter 2 talks about the overview of the project
- Chapter 3 talks about the requirement specifications for the project. Both hardware and software requirements used for the project are specified here.
- Chapter 4 provides a detailed design of the system architecture. Flowcharts and use case diagrams of the projects are used to explain the project design here.
- Chapter 5 covers the implementation aspect of the project. Here details on how the project was implemented as an end user application is specified.
- Chapter 6 covers details on the testing of the project and model built on various examples. For the model to recognise nutrient deficiency symptoms, different test images were used and the true labels of the image and predicted labels obtained from the model were recorded. Graphs showing the training and validation phase of the model are also included here.
- Chapter 7 consists the experimental results and images of the final application, with its various stages.
- Chapter 8 consists of the conclusion and future steps for the project.

The report ends with a list of references that were used in the project development

CHAPTER 2

OVERVIEW

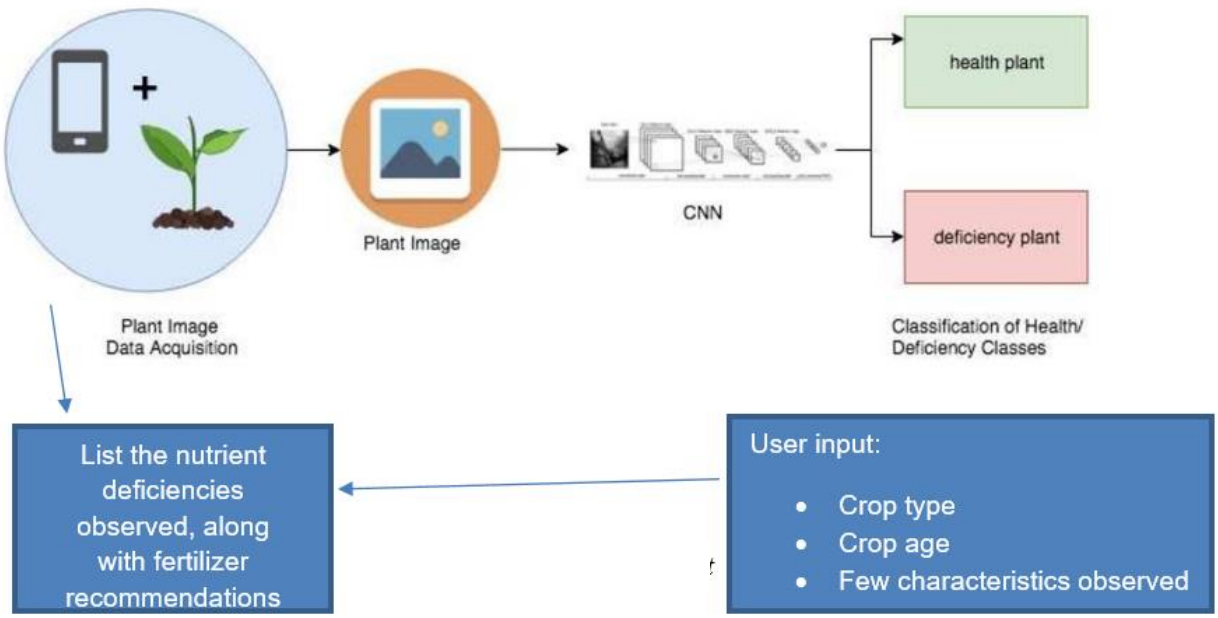

Fig 2.1 Process and flow of the project

1) Collect dataset of rice, wheat, maize leaves, and label based on symptoms
2) Build a images classification model using deep learning techniques
3) The model should take the leaf image as input and classify it based on symptoms observed
4) Create a mobile application that allows user to take a picture of the leaf and upload it to the application for analysis
5) Save and deploy the trained model to the end user application , so that it can run on the image captured by the user
6) Predict the deficiencies present based on the symptoms identified and display on screen.
7) Accordingly recommend appropriate fertilizer that can be used.
8) Take the land size details from the farmer

9) Using the stored knowledge of fertilizer requirement for corresponding nutrient deficiency and calculate the amount of fertilizer required
10) Display the recommended amount of fertilizer to the user

CHAPTER 3

REQUIREMENT SPECIFICATION

Software and Hardware requirements can be specified

3.1 Hardware Requirements:

1) Processor – INTEL/AMD
2) RAM – Min. of 4 GB
3) Hard Disk – Built-in is sufficient
4) Basic I/O Devices
5) Internet Connection broadband with a speed of 5 Mbps or higher

3.2 Software Requirements

1) OS – Linux/Windows OS
2) Modern Web Browser
3) Python Platform – Anaconda/ Jupyter/Spyder/Google Colab
4) Python dependencies- NumPy, Pandas, ScikitLearn, TensorFlow, Keras, MatPlotlib

3.3 MAPPING OF REQUIREMENTS TO THE MODULES CHOSEN

1) **Streamlit = 0.78.0**

   Streamlit is an open-source python library that makes it easy to build beautiful custom web-apps for machine learning models

2) **Numpy== 1.16.0**

   Numpy is the fundamental package for scientific computing in python. Numpy arrays facilitate advanced mathematical types of operations on large numbers of data more efficiently and with less code

3) **Scikit learn**

   Used for getting classification report for analysing model performance

4) **Matplotlib==3.2.1**

   Matplotlib is a plotting library for the python programming language. It provides an object-oriented API for embedding plots into applications using general purpose GUI toolkit.

5) **Pandas==1.1.3**

   Pandas is a python package providing fast, flexible and expressive data structures designed to make working with relational or labelled data both easy and intuitive. It is the fundamental building block for doing practical, real-world data analysis in python

6) **TensorFlow==2.3.0**

   TensorFlow is a free and open-source software library for machine learning for training and inference of deep neural networks

7) **Pillow==8.2.0**

   Pillow is an imaging library in python that adds support for opening, manipulating and saving many different image file formats

8) Openpyxl==3.0.3

This is a python library for reading and writing excel files (to access fertilizer database

CHAPTER 4

DETAILED DESIGN

4.1 ARCHITECTURAL DESIGN:

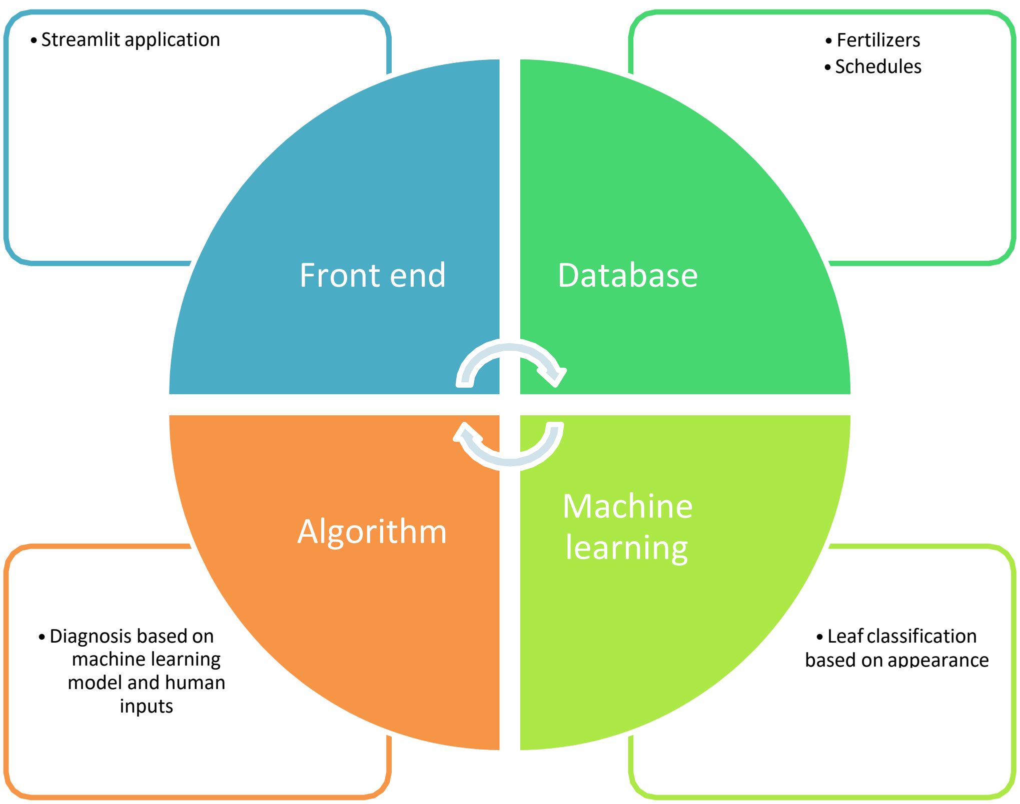

Fig 4.1.1 Architectural Design

1. **Database**

   Excel/CSV file to store Nutrient-Fertilizer data

   Data on overview of the fertilizers to use for different nutrient deficiencies

- Select land size
- Stage of growth (sapling/established/flowering)
- Crop type (rice/maize/wheat)

Default shows fertilizers recommended for all nutrients deficiencies focused on.

2. **Machine Learning**

Leaf is classified based on appearance.

3. **Algorithm:**

TensorFlow Keras Neural Network Model for Diagnosis

1) **Leaf Classification**

It helps in diagnosis of nutrient deficiencies like 'Potassium', 'Magnesium', 'Zinc', 'Iron', 'Manganese', 'Copper', 'Boron' and 'Sulphur' for rice/paddy, maize and wheat crops.

2) **Data Required**

- Image of the leaf with white back ground
- Type of crop (Rice/ Wheat/ Maize)
- Age of leaf (Mature/ Old, Young/ New, Middle)

Some human input is collected to further specialize the classification. General question about plant asked:

- Does the plant show stunted growth?
- Are there Red/dead spots on leaves?
- Are the leaves twisted/brittle?
- Is there a general yellowing of leaves observed?

3) **Neural Network Model is trained to classify the image of the leaf into the following categories:**

   - Normal
   - Spotty
   - Margin
   - Interveinal
   - Tip

4. ***User Interface***

   Stream-Lit Application

4.2 COMPONENT DESIGN

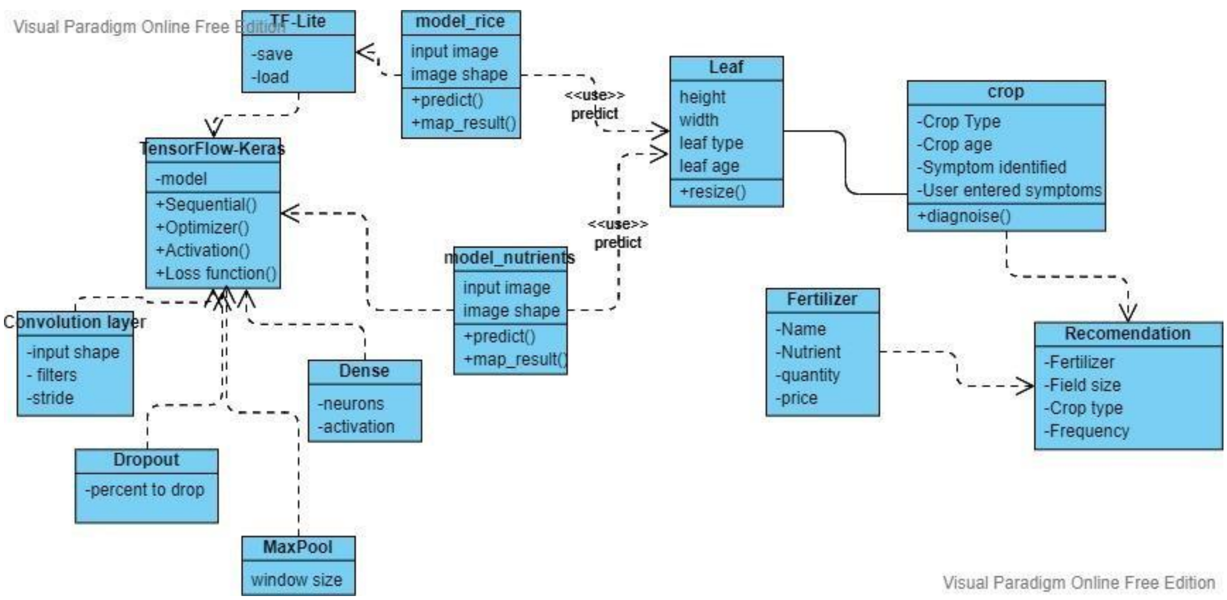

Fig 4.2.1 Component Design

4.3 BEHAVIORAL DESIGN

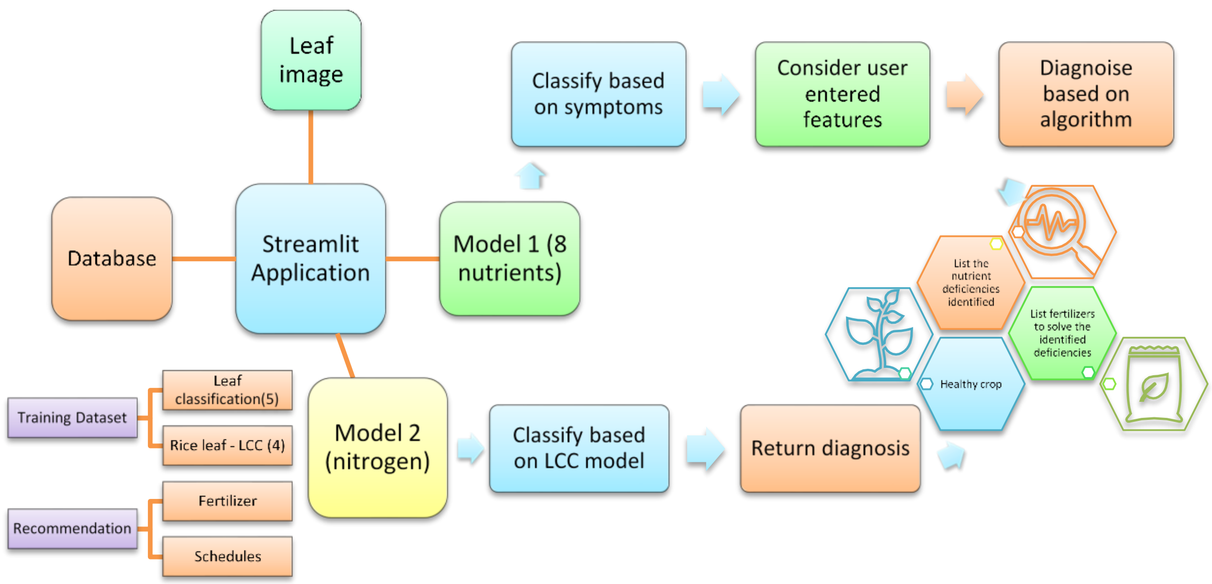

Fig 4.3.1 Behavioral Design

CHAPTER 5

IMPLEMENTATION

5.1 LEAF ANALYSIS:

The three basic tools for diagnosing nutrient deficiencies and toxicities are

1) Soil Testing
2) Plant Analysis
3) Visual observations in the field

Both Soil Testing and Plant Analysis are quantitative tests that compare soil or plant concentrations to the sufficiency range for a particular crop. Visual observation, on the other hand, is a qualitative assessment and is based on symptoms such as stunted growth or yellowing of leaves occurring as a result of nutrient stress. Interpreting visual nutrient deficiency and toxicity symptoms in plants can be difficult and plant analysis or soil testing.

Precautions in identifying nutrient stress symptoms include the following:

1) Many symptoms appear similar. For instance, N and S deficiency symptoms can be very alike, depending upon plant growth stage and severity of deficiencies.
2) Multiple deficiencies and/or toxicities can occur at the same time. More than one deficiency or toxicity can produce symptoms, or possibly an abundance of one nutrient can induce the deficiency of another. For example, excessive P causing Zn deficiency.
3) Crop species, and even some cultivars of the same species, differ in their ability to adapt to nutrient deficiencies and toxicities. For example, corn is typically more sensitive to a Zn deficiency than barley and will show Zn deficiency more clearly
4) Pseudo deficiency symptoms (Visual symptoms appearing similar to nutrient deficiency symptoms). Potential factors causing pseudo deficiency include, but are not limited to, disease, drought, excess water, genetic abnormalities, herbicide and pesticide residues, insects, and soil compaction.

5) Hidden hunger. Plants may be nutrient deficient without showing visual clues.
6) Field symptoms appear different than ‘ideal’ symptoms. Many of the were grown under controlled nutrient conditions, and deficiency/toxicity symptoms observed in the field may or may not appear as they do here. Experience and knowledge of field history are excellent aids in determining causes nutrient for stress.
7)

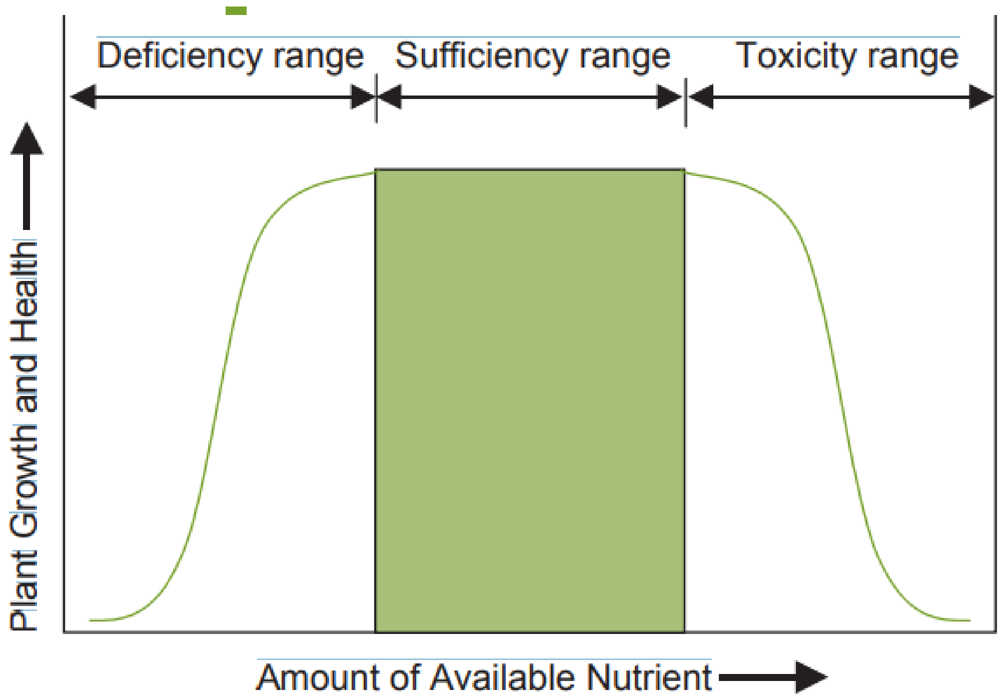

Fig 5.1.1 Plant Health versus Nutrition

In addition to the above precautions, visual observation is also limited by time. Between the time a plant is nutrient deficient (hidden hunger) and visual symptoms appear, crop health and productivity may be substantially reduced and corrective actions may or may not be effective. Therefore, regular soil or plant testing is recommended for the prevention and early diagnosis of nutrient stress. If visual symptoms are observed, record which crops are affected, their location with respect to topography, aspect, and soil conditions, a detailed description of symptoms and time of season that the symptoms first appeared. Affected field locations can be marked and monitored over time using either flagging or GPS readings. This information will be useful in preventing nutrient stress in subsequent years.

5.2 IDENTIFYING DEFICIENCY SYMPTOMS:

A first step in diagnosing nutrient deficiencies is to describe the symptoms. Each deficiency symptom is related to some function of the nutrient in the plant.

Symptoms caused by nutrient deficiencies are generally grouped into five categories:

1) Stunted growth
2) Chlorosis
3) Interveinal Chlorosis
4) Purplish-red coloring and
5) Necrosis

Stunting is a common symptom for many deficient nutrients due to their varied roles in the plant. For example, when nutrients involved in plant functions such as stem elongation, photosynthesis, and protein production are deficient, plant growth is typically slow and plants are small in stature.

Chlorosis and interveinal chlorosis are found in plants deficient of nutrients necessary for photosynthesis and/or chlorophyll (green leaf pigment involved in photosynthesis) production. Chlorosis can result in either the entire plant or leaf turning light green to yellow, or appear more localized as white or yellow spotting.

Interveinal chlorosis occurs when certain nutrients [B, Fe, magnesium (Mg), Mn, nickel (Ni) and Zn] are deficient.

Purplish-red discolorations in plant stems and leaves are due to above normal levels of anthocyanin (a purple colored pigment) that can accumulate when plant functions are disrupted or stressed. This symptom can be particularly difficult to diagnose because cool temperatures, disease, drought and even maturation of some plants can also cause anthocyanin to accumulate (Bennett, 1993). Certain plant cultivars may also exhibit this purple coloring.

Necrosis generally happens in later stages of a deficiency and causes the parts of the plant first affected by the deficiency to brown and die. Since a number of nutrient deficiencies can produce similar symptoms, further evaluation of symptoms related to particular leaf patterns or locations on the plant are needed to diagnose nutrient specific deficiencies.

Another step in identifying deficiency symptoms is to determine whether the deficiency is the result of a **mobile or immobile nutrient** based on where the symptom appears in the plant. Mobile nutrients are nutrients that are able to move out of older leaves to younger plant parts when supplies are inadequate. Mobile nutrients include N, P, K, Cl and Mg. Because these nutrients are mobile, visual deficiencies will first occur in the older or lower leaves and effects can be either localized or generalized. In contrast, immobile nutrients [B, Ca, Cu, Fe, Mn, Ni, S and Zn] cannot move from one plant part to another and deficiency symptoms will initially occur in the younger or upper leaves and be localized. Zn is a partial exception to this as it is only somewhat immobile in the plant, causing Zn deficiency symptoms to initially appear on middle leaves and then affect both older and younger leaves as the deficiency develops. It consists of alternative statements about the appearance of different plant structures. If possible, it may be helpful to have a healthy plant on hand for comparison purposes. Bear in mind that deficiency symptoms differ among crop types and plant specific amounts of N in cereals will result in few tillers, slender stalks, short heads, and grain with low protein content. Leaf curling and small tubers are common in potatoes deficient of N. Fields deficient in N can be either uniform or patchy in appearance, depending on the cause of the deficiency. For example, fertilizer application ‘misses’ will result in uniform strips of deficiency, whereas N deficiency as a result of soil characteristics such as organic matter content will be patchy. Phosphorus (P) Plants require P for the development of ATP, sugars and nucleic acids. Cool soils during the early growing season may be a factor causing P deficiency. P deficiency symptoms are usually more noticeable in young plants, which have a greater relative demand for P than more mature plants. P deficient plants generally turn dark green (both leaves and stems) and appear stunted. Older leaves are affected first and may acquire a purplish discoloration due to the accumulation of sugars which favors anthocyanin synthesis; in some cases, leaf tips will brown and die. Plants suffering from P deficiency appear weak and maturity is delayed. Small grains with P deficiency tend to be stressed and predisposed to root rot diseases, and some cultivars will turn red or purple. In alfalfa, an upward tilting of leaflets may occur. Potato P deficiency symptoms include leaves curling upward and tubers having brown internal specks, often radiating out from the core. P deficiency in corn is usually visible in young plants with leaves turning purple. From a field perspective, P deficiency is likely to occur on tops of ridges or other exposed areas that have highly eroded or weathered soils, or in areas that have been leveled, exposing subsoil high in calcium carbonate ($CaCO_3$). Crops grown in soils high in $CaCO_3$ are prone to P deficiency due to the precipitation of insoluble Ca-P minerals.

The first problem we faced was we had to identify what exact symptoms can be useful to diagnose a particular plant as having a particular deficiency. During our research, based on the research activities of Montana State University & US department of agriculture about this problem, we found that below are the symptoms which farmers can use to identify the deficiencies. Then we identified some symptoms which could be understood by the computer based on training. The key difference between pest-related problem and nutrient problem is this chlorosis issue. In pest-related problem mostly we will not notice this chlorosis issue. But in nutrient deficiency, majority of the times there will be an issue of chlorosis or change in colour and dying of cells.

Based on these specifics we have concluded symptoms to 5 classes which we will be using for training the model. Each deficiency symptom is related to some function of the nutrient in the plant.

- Tip
- Spotty
- Normal
- Margin
- Interveinal

1) Tip:

Fig 5.2.1. Image of Maize leaf with label as a ‘tip’

2) Spotty:

Fig 5.2.2. Image of wheat leaf with label as ‘spotty’

3) Normal:

Fig 5.2.3 Image of wheat leaf with label as ‘normal’

4) Margin:

Fig 5.2.4. Image of rice leaf with label as ‘margin’

5) Interveinal:

Fig 5.2.5. Image of maize leaf with label as ‘interveinal’

5.3 MODEL:

A) PROGRAMMING LANGUAGE USED: *Python*

Python is an interpreted high-level general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant indentation. We chose python over other languages because:

1) **It has great a great library ecosystem:**
A great choice of libraries is one of the main reasons python is the most popular programming language used for AI. ML requires continuous data processing, and python's libraries let you access, handle and transform data. Tensor Flow for working with deep learning by setting up, training, and utilizing artificial neural networks with massive datasets. Keras for deep learning allows fast calculations and prototyping, as it uses the GPU in addition to the CPU of the computer. Matplotlib for creating 2D plots, histograms, charts, and other forms of visualization.

2) **Flexibility:**
   - It offers an option to choose either to use OOPs or scripting
   - There's also no need to recompile the source code, developers can implement any changes and quickly see the results
   - Programmers can combine Python and other languages to reach their goals.

3) **Platform Independence:**
Python is not only comfortable to use and easy to learn but also very versatile. Python for machine learning development can run on any platform including Windows, MacOS, Linux, UNIX, and twenty-one others. To transfer the process from one platform to another, developers need to implement several small-scale changes and modify some lines of code to create an executable form of code for the chosen platform. Developers can use packages like PyInstaller to prepare their code for running on different platforms.

4) **Readability:**

Python is very easy to read so every python developer can understand the code of their peers and change, copy or share it. There's no confusion, errors or conflicting paradigms, and this leads to a more efficient exchange of algorithms, ideas, and tools between AI and ML professionals. There are also tools like IPython available, which is an interactive shell that provides extra features like testing, debugging, tab-completion, and others, and facilitates the work process.

5) **Good Visualization options:**

Python offers a variety of libraries, and some of them are great visualization tools. For Artificial Intelligence, deep learning, and machine learning, it's vital to be able to represent data in a human-readable format. Libraries like Matplotlib allow data scientists to build charts, histograms, and plots for better data comprehension, effective presentation, and visualization. Different application programming interfaces also simplify the visualization process and make it easier to create clear reports.

B) FRONT-END: *StreamLit*

StreamLit is an open-source Python library that makes it easy to create and share beautiful, custom web apps for machine learning and data science. It allows you to create a stunning looking application with only a few lines of code. A few of the advantages of using Streamlit tools like Dash and Flask:

- It embraces Python scripting; No HTML knowledge is needed
- Less code is needed to create a beautiful application
- No call-backs are needed since widgets are treated as variables
- Data caching simplifies and speeds up computation pipelines.

C) DATABASE: *csv files*

We have used Excel/CSV file to store Nutrient-Fertilizer data**.**

Data on overview of the fertilizers to use for different nutrient deficiencies**:**

- Select land size
- Stage of growth (Sapling/ Flowering/Established)
- Crop type (Rice/Maize/Wheat)

Default shows fertilizers recommended for all nutrients deficiencies focused on.

D) DATASET:

We have collected 100 images containing 5 different classes- Tip, Spotty, Normal, Margin and Interveinal of 3 crops- Rice, Wheat and Maize. The reason for selecting these three crops is they all have similar leaf structure and parallel venation.

People will not click the photo of image in a horizontal way or a vertical way every time, they randomly click, sometimes they zoom in & click, sometimes they zoom out & click. This is an edge case. So we augmented these 100 images by changing their angle & orientation and increased it to 400 images.

We then split this Dataset into 3 sets.

1) Training Data – 70%

The data used to train the model to predict the outcome we design our model to predict. A training dataset is a dataset of examples used during the learning process and is used to fit the parameters.

2) Validation Data – 20%

A validation dataset is a dataset of examples used to tune the hyper parameters. Our goal is to find the network having the best performance on new data. The simplest approach to the comparison of different networks is to evaluate the error function using data which is independent of that used for training. Various networks are trained by minimization of an appropriate error function defined with respect to a training data set. The performance of the networks is then compared by evaluating the error function using an independent validation set, and the network having the smallest error with respect to the validation set is selected.

3) **Test Data – 10%**

A test dataset is a dataset that is independent of the training dataset, but that follows the same probability distribution as the training dataset. If a model fit to the training dataset also fits the test dataset well, minimal over-fitting has taken place. A better fitting of the training dataset as opposed to the test dataset usually leads to over-fitting. Test data is used to measure the performance, such as accuracy or efficiency, of the model developed.

5.4 MODEL STRUCTURE:

TensorFlow-Keras Neural Network model for Diagnosis.

Data required:

1) Image of the leaf with white background
2) Type of crop (Rice/Wheat/Maize)
3) Age of leaf (Sapling/Flowering/Established)

Some human input is collected to further specialize the classification. General questions about plant:

a) Does the plant show stunted growth?
b) Are there red/dead spots on leaves?
c) Are the leaves twisted/brittle?
d) Is there a general yellowing of leaves?

Neural Network Model is trained to classify the image of the leaf into the following categories:

- Normal
- Spotty
- Margin
- Interveinal
- Tip

Preprocesing – Rescale image pixel values

- Convolution layer – 32 filters, stride 3, activation = ReLU

MaxPool layer

- Convolution layer – 64 filters, stride 5, activation = ReLU

MaxPool layer

- Dropout Layer (drop 50% of neurons)
- Convolution layer – 64 filters, stride 5, activation = ReLU

MaxPool layer

- Dropout Layer (drop 50% of neurons)
- Convolution layer – 64 filters, stride 5, activation = ReLU

MaxPool layer

- Dropout Layer (drop 40% of neurons)
- Dense layer – 128 neurons, activation = ReLU
- Dense layer = 5 neurons, activation = softmax

Preprocessing refers to all the transformations on the raw data before it is fed to the machine learning or deep learning algorithm. Training a convolutional neural network on raw images will probably lead to bad classification performances. So we have normalized by dividing with 255, which is the RGB range.

Convolution layer is to identify the features in images, where we multiply the matrix with the image matrix of which the end result will be the smaller size of the image.

Rectified Linear Unit activation, ReLU is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. It is the default activation function for many types of neural networks because a model that uses it is easier to train and often achieves better performance.

MaxPool Layer helps reduce the spatial size of the convolved features and also helps reduce over-fitting by providing an abstracted representation of them. It is a sample-based discretization process. It is similar to the convolution layer but instead of taking a dot product between the input and the kernel we take the max of the region from the input overlapped by the kernel. Also we don't need the values of background but only the numbers where it's changing, i.e., where exactly the color is changing. The MaxPool Layer replaces the whole box by 3*3. It's a way of reducing the dataset, like compressing so that it’ll be faster to compute the next layer.

Drop-Out Layer is a simple way to prevent neural networks from OverFitting. Drop out layer randomly sets input units to 0 with a frequency of rate at each step during training time, which helps prevent OverFitting. Inputs not set to 0 are scaled up by 1/(1 - rate) such that the sum over all inputs is unchanged. This layer only applies when training is set to “True” such that no values are dropped during inference. When we tried without dropout layer also & what we found that it was OverFitting the training data but after using this drop out layer we found it was not OverFitting much, the validation score was little higher.

OPTIMIZATION:

Stochastic Gradient Descent is an optimization algorithm that can be used to train neural network models. The Stochastic Gradient Descent algorithm requires gradients to be calculated for each variable in the model so that new values for the variables can be calculated. It picks one data point from the whole dataset at each iteration to reduce the computation enormously. i.e., Instead of changing weight for every example, we run for all the examples & after that we change.

LOSS FUNCTION:

Categorical cross entropy is a loss function that is used in multi-class classification tasks. These are tasks where an example can only belong to one out of many possible categories, and the model must decide which one. We have used this function because our model is a multi-class model.

CHAPTER 6

TESTING

An epoch means training the neural network with all the training data for one cycle. In an epoch, we use all of the data exactly once. A forward pass and a backward pass together are counted as one pass: An epoch is made up of one or more batches, where we use a part of the dataset to train the neural network

```
plt.plot(history_.history['accuracy'], 'b')
plt.plot(history_.history['val_accuracy'], 'g')
plt.title('model accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')
plt.show()

```

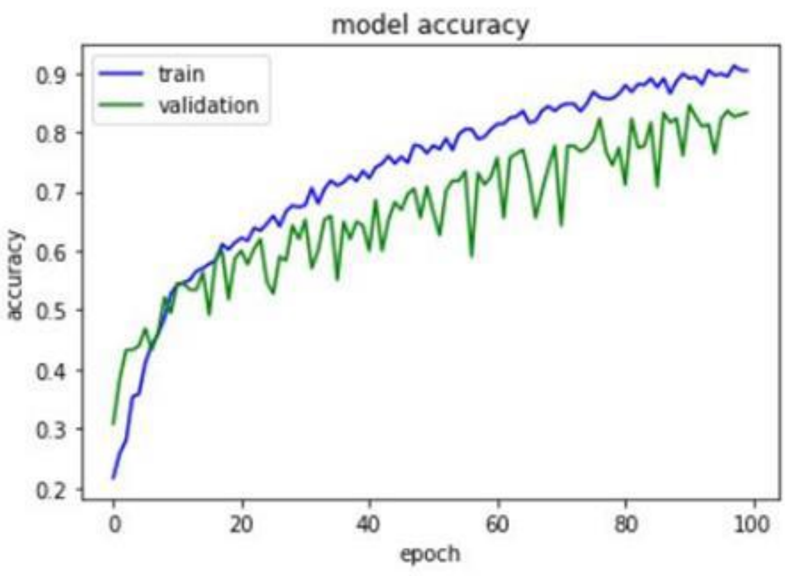

Fig 6.1. Increase in accuracy when training for 100 epochs

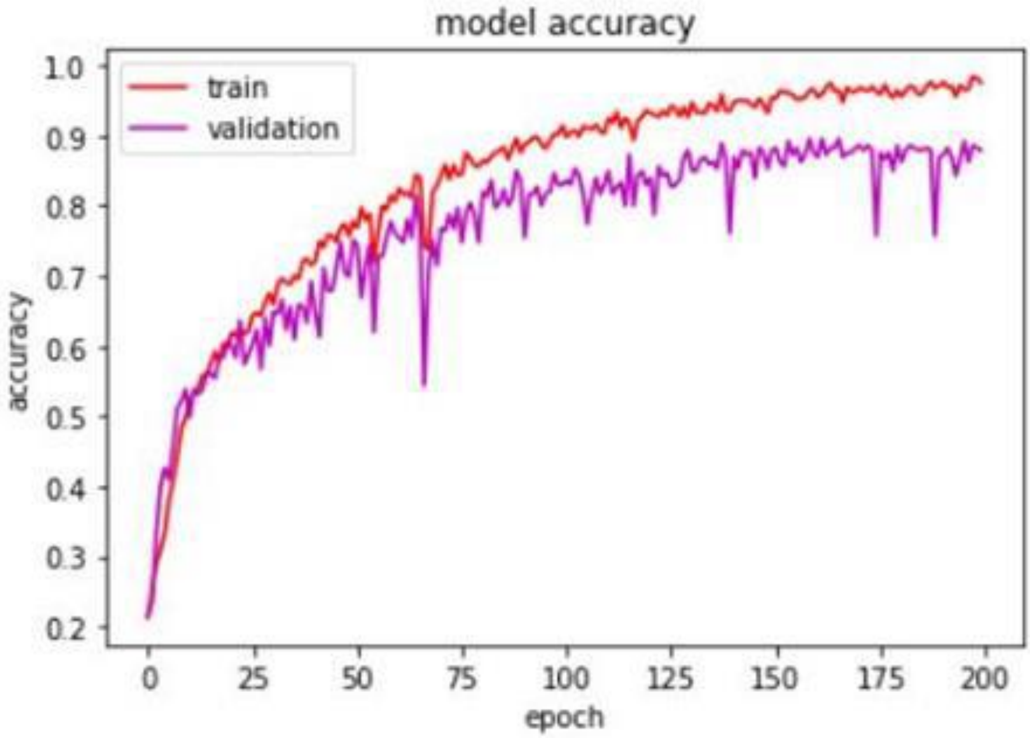

Fig 6.2. Increase in accuracy when training for 200 epochs

```
# summarize history for loss
plt.plot(history_.history['loss'], 'b')
plt.plot(history_.history['val_loss'], 'g')
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')
plt.show()
```

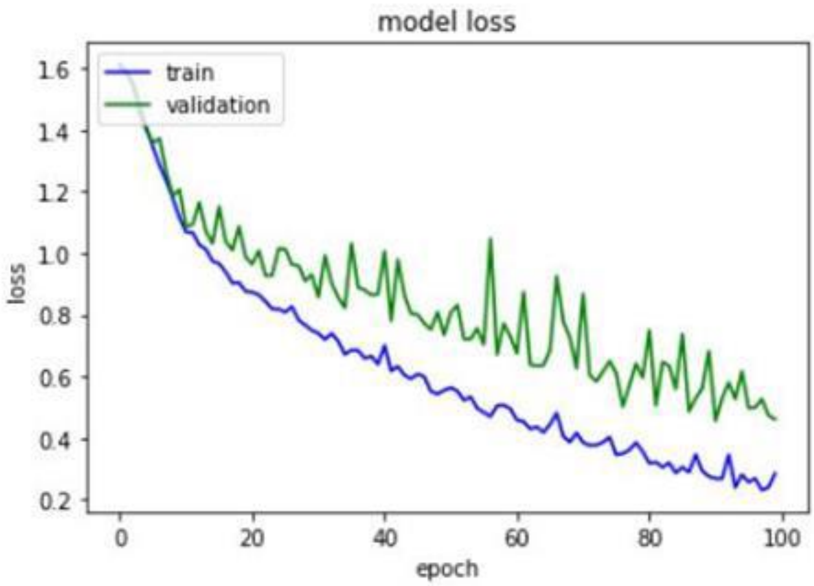

Fig 6.3. Decrease in loss when training for 100 epochs

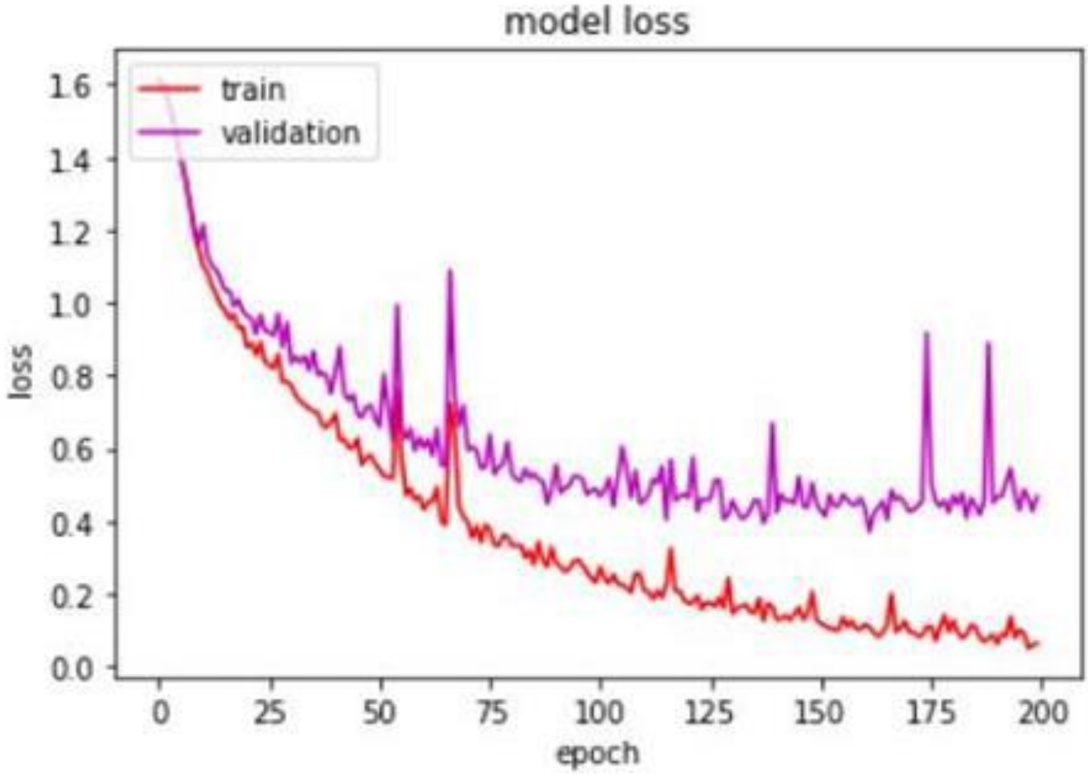

Fig 6.4. Decrease in loss when training for 200 epochs

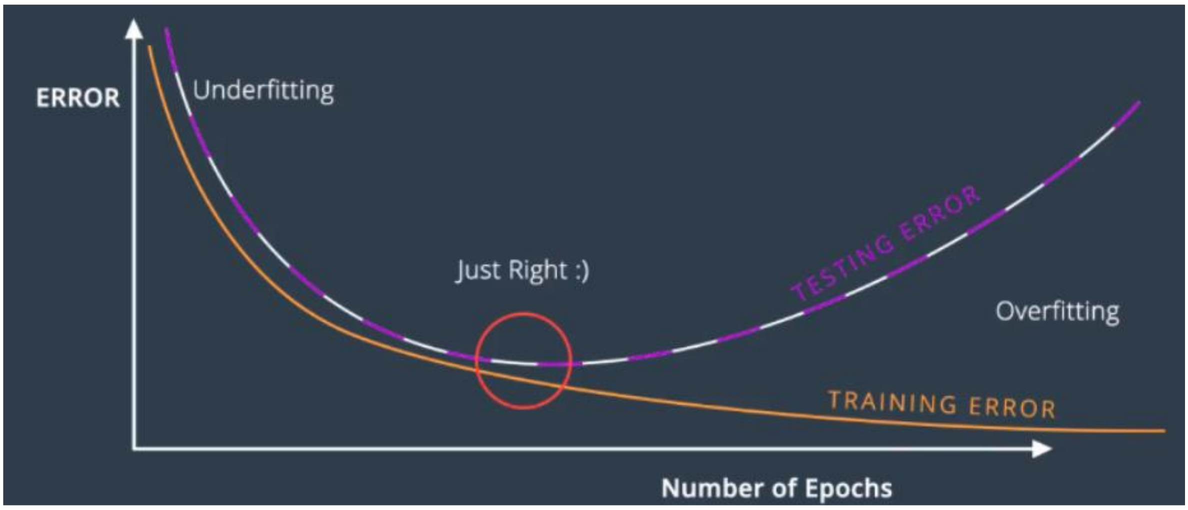

Fig 6.5. Number of Epochs versus Error

The number of epoch will decide how many times we will change the weights of the networks. As the number of epoch increases, the same number of weights are changed in the neural network and the boundary goes from under-fitting to over-fitting

As shown in the above figures, we have trained to the model two times. Once for 100 epochs in which we found that the model is slightly under-fitted on the data due to which the accuracy is slightly lower. Then for 200 epochs where found the model fitted on the data perfectly due to which the accuracy is higher than the model for 100 epochs. So we chose the model for 200 epochs because of its better accuracy.

Finally we got a **Train, Validation accuracy of 90%** and **Final Test accuracy of 88.24%.**

6.1 MODEL PREDICTIONS EXAMPLES

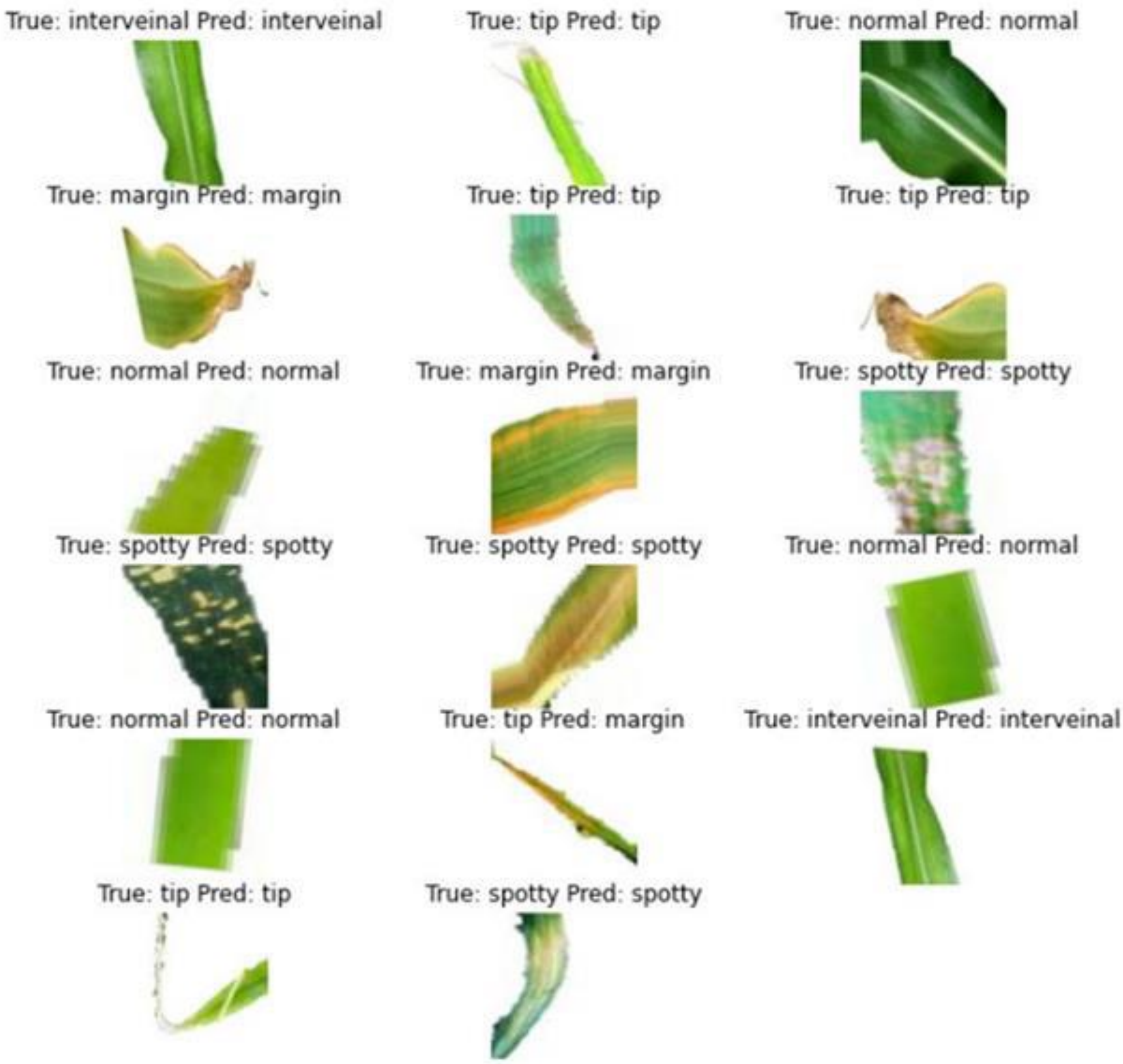

Fig 6.6. Examples of model predictions along with True labels

6.2 CLASSIFICATION REPORT:

```
Confusion Matrix
[[41  0  0  0  0]
 [ 0 33  0  4  4]
 [ 0  0 41  0  0]
 [ 2  0  0 38  1]
 [ 0  6  3  0 32]]
Classification Report
              precision    recall  f1-score   support

  interveinal       0.95      1.00      0.98        41
       margin       0.85      0.80      0.83        41
       normal       0.93      1.00      0.96        41
       spotty       0.90      0.93      0.92        41
          tip       0.86      0.78      0.82        41

     accuracy                           0.90       205
    macro avg       0.90      0.90      0.90       205
 weighted avg       0.90      0.90      0.90       205
```

Fig 6.7. Final Model Confusion Matrix & Classification Report

The classification matrix is a standard tool for evaluation of statistical models and is sometimes referred to as a confusion matrix. A classification matrix is an important tool for assessing the results of prediction because it makes it easy to understand and account for the effects of wrong predictions. From classification report, we observe that the model is slightly confused between tip and spotty classes. However, the model is showing high accuracy for the remaining classes.

CHAPTER 7

RESULTS

7.1 LEAF DIAGNOSIS:

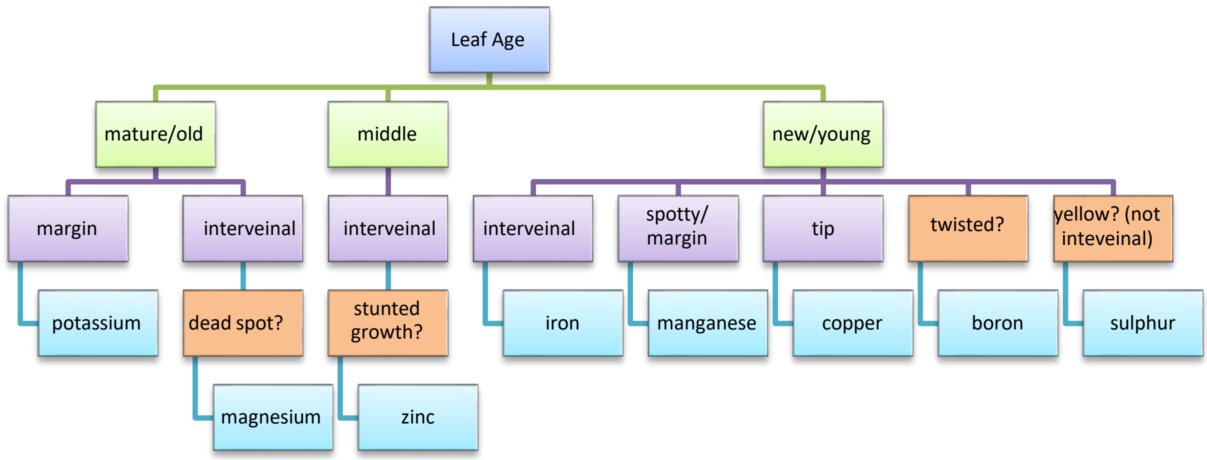

Fig 7.1.1 Diagnosing Nutrient deficiency

Green – User Input

Purple – 5 classes

Orange – Additional User Input

Blue – Identified Nutrient Deficiency

First, we have leaf age, which is the user input. When user uploads the plant image and enters its age & crop type, the model will classify it into one of the five classes – Margin, Interveinal, Spotty, Normal, and Tip. Once the image is classified into one of these five, based on the additional user information, like if the leaf is twisted or have a dead spot or if it's showing yellowing or stunted growth, the model identifies which nutrient is deficient and use it to recommend the type, quality and quantity of fertilizers to apply to the fields.

7.2 APPLICATION FLOW:

Fig 7.2.1 Chart showing the Application Flow

7.3 SCREENSHOTS:

7.3.1 Home Page:

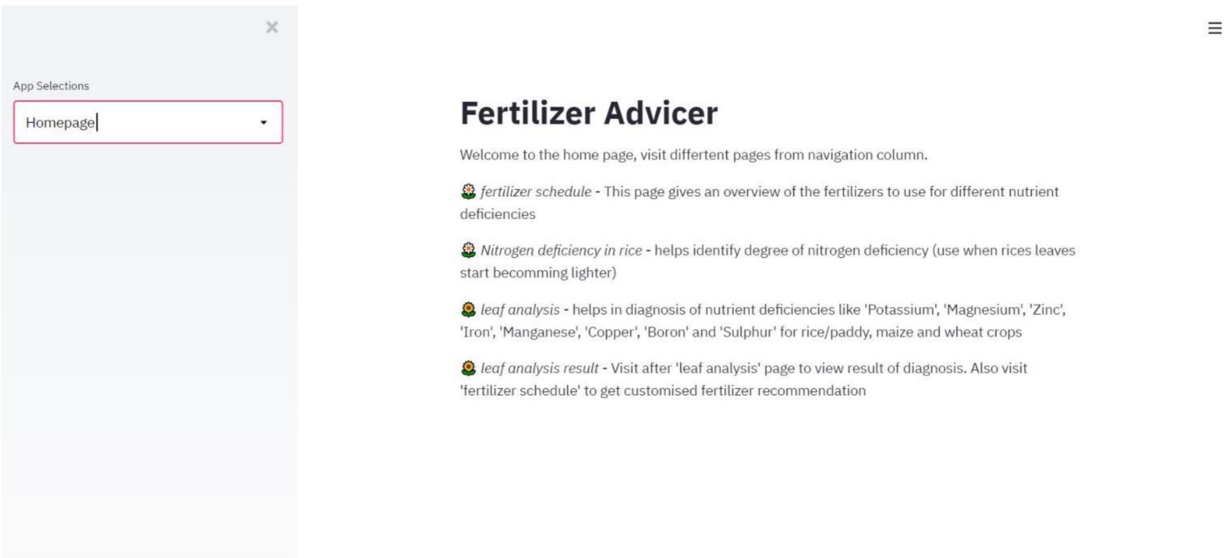

Fig 7.3.1 Home Page

Home page gives the description of the application and also navigation to individual pages Nutrient Deficiency Detection in Plants for Fertilizer Management

7.3.2 Leaf Analysis Page:

Fig 7.3.2 Leaf Analysis Page

Leaf Analysis helps in diagnosis of ***8 nutrient deficiencies like 'Potassium', 'Magnesium', 'Zinc', 'Iron', 'Manganese', 'Copper', 'Boron' and 'Sulphur'*** for ***3 crops Rice/Paddy, Maize and Wheat crops***

- Choose age of leaf (Mature/Old, New/Young and Middle)
- Choose crop (Rice/Maize/Wheat)
- Upload image of leaf with white background
- Click button to get diagnosis

The Neural Network model trained to classify an image into the five categories- Tip, Spotty, Normal, Margin and interveinal is run. Here some human input is collected to further specialize the classification. General question about plant asked:

- Does the plant show stunted growth?
- Are there Red/dead spots on leaves?
- Are the leaves twisted/brittle?
- Is there a general yellowing of leaves observed?

Along with answers to these questions and the results of the model, nutrient deficiency if identified is displayed in this page.

7.3.3 Fertilizer Schedule Page:

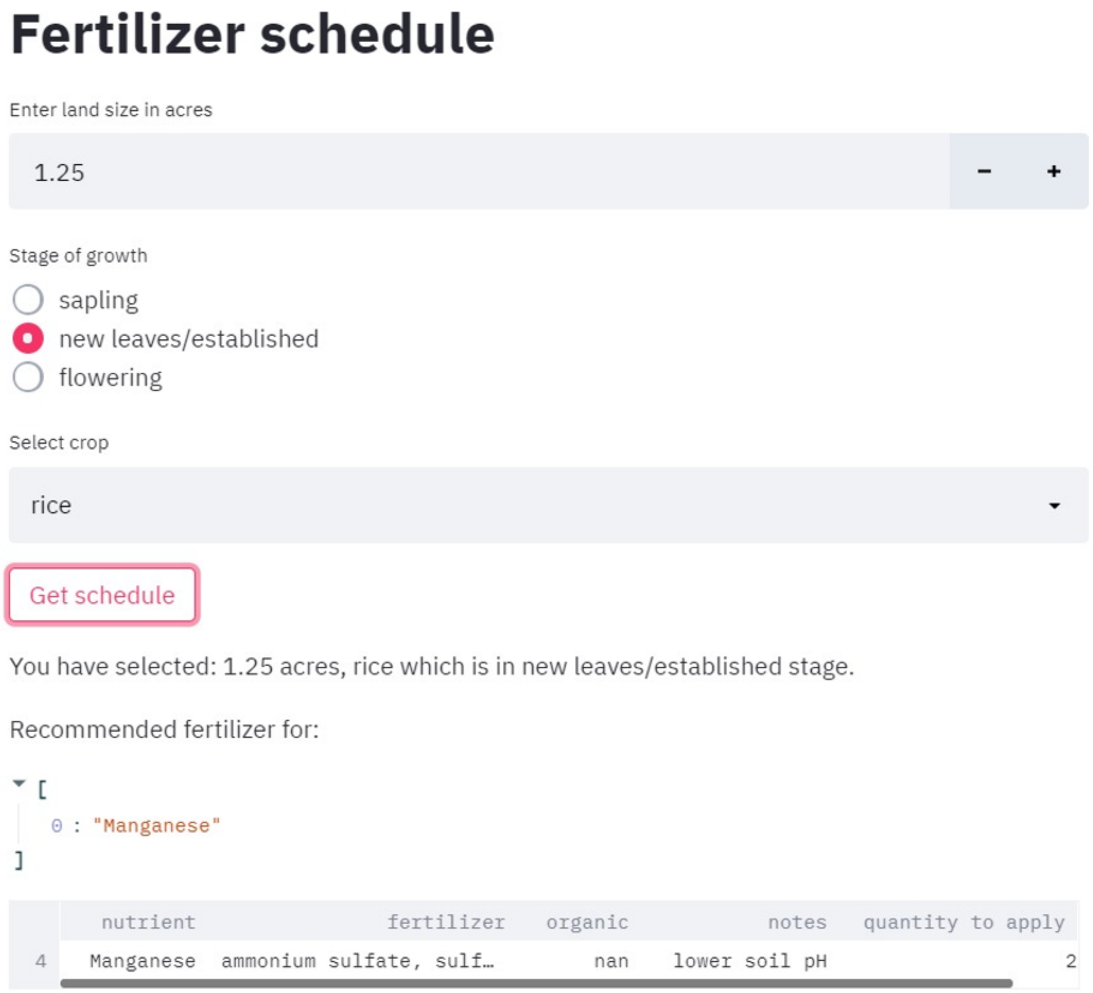

Fig 7.3.3 Fertilizer Schedule Page

This page gives an overview of the fertilizers to use for different nutrient deficiencies

- Select land size
- Stage of growth (sapling/established/flowering)
- Crop type (rice/maize/wheat)

Default shows fertilizers recommended for all nutrients deficiencies focused on.

7.3.4 Results Page:

Fig 7.3.4 Results Page

Result shows the complete diagnosis of the plant. It shows nutrient deficiencies, if identified, along with the confidence value.

CHAPTER 8

CONCLUSION

We have used computational tools for nutrition monitoring as a part of decision support & farm management tools, to develop this model which is of immense use to the farmers who do not have access to expert advice. We have developed an application that uses the leaf images of crops like rice, wheat and maize to identify nutrient deficiencies like Potassium, Magnesium, Zinc, Iron, Manganese, Copper, Boron, and Sulphur along with some additional human input and recommends appropriate type and amount of fertilizer.

We have equipped a machine learning model to classify the unhealthy leaves so that using this application through mobile or web, farmers can make informed decision to buy the adequate proportion of the fertilizer to treat their plants with the correct nutrients.

We are planning to make this project a little more useful and reachable to our farmers with our future enhancements

Scope for further work:

- To expand this project to other vegetable and fruits crops.
- To save farmer's data regarding the crop, area etc., and use it to predict what fertilizer is needed, in advance, before the symptoms appear.
- To also use this farmer's data to recommend shops to sell produce.

REFERENCES

1. G. Xu, F. Zhang, S. G. Shah, Y. Ye, and H. Mao, “Use of leaf colour images to identify nitrogen and potassium deficient tomatoes,” Pattern Recognition Letters, vol. 32, no. 11, pp. 1584–1590, Aug. 2011.

2. K. A. M. Han and U. Watchareeruetai, "Classification of Nutrient Deficiency in Black Gram Using Deep Convolutional Neural Networks," 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE), Chonburi, Thailand, 2019, pp. 277- 282, doi: 10.1109/JCSSE.2019.8864224.

3. C. A. Dhawale, S. Misra, S. Thakur and N. Dattatraya Jambhekar, "Analysis of nutritional deficiency in citrus species tree leaf using image processing," 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, 2016, pp. 2248-2252, doi: 10.1109/ICACCI.2016.7732386.

4. U. Watchareeruetai, P. Noinongyao, C. Wattanapaiboonsuk, P. Khantiviriya and S. Duangsrisai, "Identification of Plant Nutrient Deficiencies Using Convolutional Neural Networks," 2018 International Electrical Engineering Congress (iEECON), Krabi, Thailand, 2018, pp. 1-4, doi: 10.1109/IEECON.2018.8712217.

5. M. Latte, S. Shidnal, and B. Anami, “Rule Based Approach to Determine Nutrient Deficiency in Paddy Leaf Images,” International Journal of Agricultural Technology, vol. 13, no. 2, pp. 227–245, 2017.

6. S. Sivagami and S. Mohanapriya, “Automatic Detection of Tomato Leaf Deficiency and its Result of Disease Occurrence through Image Processing,” 2019. [Online]. Available: https://www.semanticscholar.org/paper/Automatic-Detection-of-Tomato-LeafDeficiency-and-Sivagami-Mohanapriya/e06b426a70ed1b1cf2f69dad7aaf0813ddcdf070. [Accessed: 28-Sep-2020].