NOTE: ALL THE POINTS WRITTEN ARE OPEN FOR DISCUSSION

**Brief Intro:**

We have 2 parts of our problem statement, due to the fact that making a solution for both of these parts will be a hectic job and we will not be able to deliver a quality and impressive solution as well if we pursued creating solution for both parts.

Therefore, we have decided to create solution for only [**First part of our PS**](#def_1).

With that in mind, lets move ahead:

**Path Ahead:**

1. We will be working on “Disease identification and solution providing” part for now, since it is the main subset of our overall [**First Part of PS**](#def_1)**.**
2. Our aim should be to deliver a quality prototype, for that we have to offer features like regional language, easy to use UI, it should be intuitive for farmer to use our application i.e. he should just click photo and we will tell what disease crop is affected with and what are solutions.
3. Currently following datasets for disease identification have been found:

* Plant Village
* Plant Doc

Plant Village dataset is a large one and is better, the only issue is: it primarily have the images of diseased leaves from tomato, potato and bell pepper, i.e. we will have a hard time identifying disease in plants of say brinjal, garlic, etc.

1. In parallel we will keep having discussion about crop recommendation system for farmer and if we are able to make a quality working prototype for disease identification on time then we will be working on developing this crop recommendation feature as well.

**What will be the benefit?**

* As is suggested by Dr. Kalapriya Mam, it would be better for us if we follow building the solution of [**First Part of our PS**](#def_1)**,** she mainly liked the idea of recommending crops to farmer, according to her this will be actually impactful on ground and solves real world problem.
* On adding this Crop Recommendation feature along with our Disease identification feature we will be building a strong product that will be having more appeal to judges.

**Strategy to build Disease Identification Application:**

1. We will be building a **website** instead of mobile application, following are the reasons:

* Judges don’t need to install anything. They simply click a URL and see the product, immediate accessibility.
* Building website based prototype will be easier for us as compared to building a mobile app based prototype, mobile apps can face “APK won’t install” or “iOS build fails” type of errors, and we have to keep in mind the fact that different android versions might or might not support our mobile app.
* It is easier to build a scalable product with websites.
* Judges values substance over form factor.

1. We will identify diseases the plant is facing via analysing image of leaves only, this is done to make our work easier, this will also enable us to deliver a more quality and powerful product.

Diseases can also be identified by analysing roots, stem and other parts of plants.

What leaf analysis can identify:

* Fungal Diseases
* Bacterial Diseases
* Viral Diseases
* Nutrient Deficiencies
* Early stage detection of several diseases is also facilitated by analysis of leaves.

Limitations of Leaf Analysis:

* Root and Vascular Diseases
* Latent or systematic infections
* Abiotic Stress

But But But… majority of plant diseases (we are talking about 70%-90% of plant diseases) can be identified by analysing leaves of plants.

We can keep “identifying diseases via other parts of plant” feature in our future prospect, this is a thing that our project will be scalable to.

1. Training will be done on Plant Village dataset, since it have way more number of leaf images than PlantDoc dataset, but following are the issues with Plant Village dataset:
2. This dataset primarily consist images of diseased leaves of tomato, potato and Bell Pepper plants, even if we trained a great ML model on this dataset we will only be able to identify diseases with tomato, potato and Bell Pepper crops only and this is an issue.
3. This dataset primarily consists leaves effected with diseases and have almost no image of leaves suffering from pest damage or nutrient deficiency.
4. Suppose we have made a model that is able to identify disease as well nutrient deficiency, still there exists a scenario where a leaf is suffering from both nutrient deficiency and a fungal disease at the same time, the issue is: for our ML model which is trained on leaves with unique problem (i.e. leaf is either having nutrient deficiency or a disease) the possibility where both problems exists at the same time simply doesn’t exists.
5. Solutions of problems with Plant Village dataset:

Solution of problem 1:

1. We can do as our mentor Dr. Kalapriya Mam suggested, train ML model to identify diseases in approx. 5-6 types of crops.

Issue with this approach is:

* Farmer will be able to identify diseases only with few type of crops, not practical, if our main focus is building Disease identification system we should be providing a great and practical solution that actually solves problem of farmers on ground.

1. We can work on building a GAN (Generative Adversarial Network) that will be able to translate images of leaves of other plants (say brinjal) to leaves of plants that out ML model is trained on (say leaves of tomato).

This will be specially beneficial for us since with this we will be able to identify diseases in way greater number of crops, making our feature more useful for farmers on ground and with this we will be solving a real-world problem.

Solution of problem 2:

* We can use OLID I dataset, this dataset specifically consists of nutrient deficiency and pest damage on leaves, multiple crops are used as well, have around 4500 images (plant village have 50K+)
* We can skip detecting leaves suffering from nutrient deficiency and pest damage to make our job easier, maybe we can keep this thing as future prospect (Discussion required) or if decides to build it we should market it highly as this would make our ml model way more intelligent and better than our counterparts (ideally speaking).

Solution of problem 3:

* I don’t know, you guys discuss.

1. We are yet to decide Once a disease is identified by our ML model then how are we going to suggest solutions to farmer i.e. steps farmer should follow to cure his crops from the identified disease.

Following are solutions that came to mind (open for discussion):

1. Build a web scrapper that will search internet for cures of identified disease
2. Build an ML model that is able to tell what steps should be followed to cure a certain disease.

Once this all is build successfully and we achieves desired results from our web based Disease identification system, we can start working on creating “Crop recommendation system” to deliver a full holistic product that actually solves several real world problems faced by farmers, meanwhile discussions shall proceed on both with primary focus being on disease identification.

**Technical Details and Approach:**

Coming up….

**Glossary:**

* **First Part of PS:**

We have to make an application that help farmers by recommending what to grow based on past trends of prices of different crops in different seasons, what crop can be sold for more price in coming season, etc. After suggesting farmer what to sow we will help him until harvesting of the crop is done, to do this we will help him by suggesting what is appropriate amount of water to be given to the plants, disease detection via uploading image of diseased part of plant and things like that. Weather updates and suggestions for farmers on what to do when weather is bad.  
End goal is to help farmer decide what is the best crop he can grow and increase farmer’s earnings.