PROBLEM STATEMENT

DEEP LEARNING OF PLANT TOPOGRAPHY FOR PREDICTIVE AND PRESCRIPTIVE ANALYSIS AGAINST PEST AND DISEASES

INTRODUCTION

Crop losses owing to pest & diseases are inherent in Indian agriculture with the annual loss of 15-25% of productivity. Pest & diseases are complex, crop/region specific, seasonal, epidemic/endemic, which require integrated approaches to manage the loss. Due to the level of complexity, diagnosis for preventive measures are challenging, particularly our inability to see the pest/disease occurrence and their life cycle with naked eyes, while the level of difficulty raises with the size of land holdings. Due to the poor visibility of pest and disease occurrences, our ability to integrate and use the data for preventive/prescriptive measures has been the challenge leading to continuous productivity loss.

Crop loss is generally diagnosed by the symptoms developed by pest and diseases, while manual diagnosis has a very limited scope in identifying the damaged plant parts, while recognizing the pattern of pest/disease distribution almost close to impossible in large area. Scanning the large area using both far by and nearby analysis, could help track changes in plants at large scale and identify the distribution pattern very precisely. Further, an artificial intelligence driven software could process the images; integrate the existing knowledge and develop the solutions in real-time. In addition, such approach could also open the scope of predictive analysis to execute automate alert to the farmers, potentially improving the crop productivity by more than 20%.

OUR SOLUTION

Our Idea is to design and implement an AI solution for plant disease detection in large Indian Farms where disease detection at the last stages leads to crop produce failure and farmer suicides. A predictive Deep Learning model trained on Kaggle dataset for ultra fast training will enable us to develop an extremely efficient, reliable and cost effective solution for plant disease detection using drone based/satellite imagery of the large farms. We will use a **25-layer Sequential CNN model** to classify the images into **healthy/ diseased categories**. The dataset on which the model will be trained is obtained from **Kaggle** and consists of multiple **high resolution** images belonging to the above mentioned categories. This project would indeed facilitate the farmers to identify the percentage by which the crops have been affected by pests and in accordance to the percentage they can implement some of the solutions suggested by our web application for preventing the disease to spread in the farm and thereby improving the crop yield.

IMPLEMENTATION USING CLOUD COMPUTING & CNN MODEL

1. The input image will first be preprocessed using **OpenCV** and **ScikitLearn** Python Libraries.
2. We will then be applying multiple **feature detectors** to the input image inorder to generate multiple **Feature Maps**. These Feature Maps would form the first **Convolutional Layer**.
3. To this Convolution Layer, we will be applying **Rectified Linear Unit Function** to increase **Non-Linearity** in our CNN Model to reduce **overfitting**.
4. We will then be applying **Max Pooling** to this convolution layer to get a set of **Pooled Feature Maps** which would make up the **Pooled Layer**.
5. Then there would be multiple such **Convolutional** and **Pooled Layer**s one after the other as **hidden layers**.
6. We will then apply the **Flatten Function** to convert the **Pooled Feature Maps** into **numpy vectors** that can be accepted by ANN.
7. These will then act as the **input layer** of the ANN model which would consist of **fully connected hidden layers**.
8. Inorder to deal with **Bias-Variance tradeoff**, we will implement **k-fold cross validation** to get the **variance** in the accuracy scores of k number of trainings.
9. Inorder to reduce overfitting, we will use **Dropout Regularization** to randomly deactivate neurons in each hidden layer to reduce inter layer over dependencies.
10. For **Hyperparameter Tuning**, we will use **Grid-Search Cross Validation** to predict the best set of **hyperparameters** which lead to max accuracy.
11. Because of such a deep classification model, we need to have **excellent computational power** for training such a deep neural network on a very huge dataset consisting of high resolution images.
12. This is where cloud computation will come into the picture which will allow us to do three main things: 1. Store the **Dataset** on **Cloud** and import it **directly**. 2. Train the CNN Model using **GPU** in the **Cloud**. 3. Store the resultant **h5 files** on cloud. 4. Run the **prediction algorithms** directly onto these stored output h5 files.

PROJECT DELIVERABLES

1. A Managed communication of the images/data to the cloud for **real time computation**.
2. AI driven Software – fully functional for predictive and prescriptive analytics
3. **Alert system** for users (including farmers) in various formats through our web application.
4. **Multiple language support** powered by Google Translator for a user friendly interface for our web application .

RESULTS

Our app will provide the following results :

1. Predict the disease based on the image of the plant leaf with an **accuracy of 99%**.
2. Prescribe the **best possible solution** for the predicted disease from the database.

SHOWSTOPPER

- The Disease detection algorithm will be applied parallelly to multiple images of the same farm and the results will be combined to get the most precise output.
- **Parallel analysis** will lead to drastic decrease in time for the analysis.
- The result of this analysis will then be displayed in an android based smartphone with the help of a web-app, which will have a very interactive UI that is simple to understand by the end user, which in our case is a farmer .
- The farmers who use our technology will be trained on how to use the mobile application.
- As we need to retrieve and store large amount of complex data, **cloud-native computing** would help us to accommodate the storage and computing needs for faster computation
- **Data lakes** housed in the cloud access to bigger, better data for training without straining any of our in-house resources. This would help in deploying newer models for continuous innovation and for better strategic advantage.
- Along with the disease detection, the temperature, humidity and other physical parameters of the environment will be computed with the help of sensors mounted on our drone. Based upon the collective output of all the sensors and the machine learning algorithm, the farmer will be advised as to what are the best possible solutions to fight with the disease and maximize their crop yield.
- These solutions will be based upon the known approaches to tackle the diseases in various stages and variable physical conditions and will be stored in a database capable of operating in offline mode.
- By using cloud computing in our proposed solution we can package our application‘s into a container, allowing us to **replicate the results across multiple platforms**, so that the performance of our solution is **platform-independent** thus ensuring a robust digital transformation framework .
- Also, the feedback of the effectiveness of the solution advised to the farmer will update the database of solutions, so that our technology is up to date with the changes in the pattern of spreading of diseases in different conditions.

WHAT NEXT

We have presently developed a prototype trained on a relatively small dataset consisting of just some plants. Moreover, our application is also not supported on IOS

Devices. We wish to **scale up our prototype** into a market level product by inculcating **multi device and cross platform support** in our application and also training our model on much bigger datasets with very high accuracy and cover a large set of plant diseases which not only occur in India, but also in the entire world at large.

We also wish to keep working on this project as part of our final year major project thereby scaling it to a highly practical level.We plan to further develop our idea by converting our web application into a mobile app so that it would be more convenient to use and we would strive to use drones for taking images of large fields where it is not feasible to manually capture the images.

We also plan to test it on the Indian Farms specially in the Vidarbha Region where farmer suicides is an urgent issue which need to be tackled by bridging the gap between technological advancement and agriculture.