

---

# Plant Dr

— Plant Identification and diagnosis  
machine learning model —

---

# Overview

Goal: Develop a machine learning model capable of:

- identifying different plant species based on common names (Tomato, Mango, Petunia, etc.)
- diagnosing various plant ailments such as:
  - Blight
  - Rust
  - Powdery Mildew

Idea behind it:

- It is difficult not only to distinguish between plant species due to similarity in characteristics, but also plant ailments due to the similar nature of their symptoms.
- This model aims to learn and utilize relevant plant images and information to act as a trained “doctor” to take care of plant identification and diagnosis for you.

# Industry Relevance

## Agricultural industry



Farmers



Nurseries



Researchers

## Home gardening



Retail Stores



Hobbyists

# Data Transformation

- Original 8 different datasets:
  - 102 Oxford Flowers
  - Collection of Different Category of Leaf Images
  - Corn Leaf Diseases
  - Banana Leaf Nutrient Deficiencies
  - Major Crop Leaf Diseases
  - Open Leaf Image Dataset (OLID) of Bangladesh's Major Crops
  - Common plant image dataset
  - Dataset for Crop Pest and Disease Detection
- All were formatted differently in their folder structure

# Data Transformation

- Re-labeling of images
- Merging of 4 datasets into one master dataset
- Creation of a csv file for image labeling
- Image augmentation
  - `rotation_range=20`: Rotates images randomly 20 degrees.
  - `width_shift_range=0.2`: Shifts images horizontally by 20% of the width.
  - `height_shift_range=0.2`: Shifts images vertically by 20% of the height.
  - `shear_range=0.2`: Applies a shear transformation to the images.
  - `zoom_range=0.2`: Zooms into images randomly by up to 20%.
  - `horizontal_flip=True`: Randomly flips images horizontally.
  - `fill_mode='nearest'`: Uses the nearest pixel value to fill in newly created pixels after transformation.

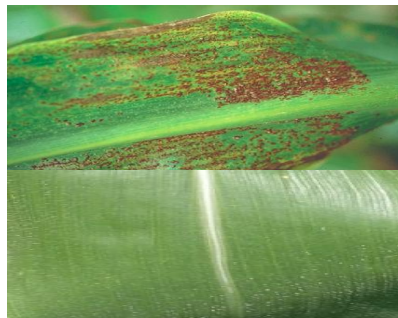
# Master Dataset



- 102 Oxford Flowers
- Compiled at the University of Oxford
  - 102 species of flowers found in the UK
  - 6,553 images



- Leaf images
- Compiled at Shri Mata Vaishno Devi University
  - 11 plant species in healthy and diseased states.
  - 4,503 images



- Corn diseases
- Combination of two other datasets ("PlantVillage" and "PlantDoc")
  - 4 states of health
  - 4,189 images



- Major Crop Diseases
- First published in the "Computers & Electrical Engineering" journal on ScienceDirect
  - 14 plant species
  - 18 different states of health
  - 61,487 images

Total number of images:

76,731 (531.5 MB)

# ML Models

## - VGG16

- pre-trained convolutional neural network (CNN) architecture
- widely used for image classification tasks.
- Pros:
  - effective for image classification tasks when you have a large, labeled dataset
- Cons:
  - slower to train and deploy
- **Accuracy Score: 72.37%**

## - EfficientNetB0

- EfficientNet is a family of models that scale efficiently in terms of depth, width, and resolution, with EfficientNetB0 being the smallest model in the family.
- Pros:
  - better performance with fewer parameters and computations
- Cons:
  - architecture is more complex, making it harder to understand and implement from scratch.
- **Accuracy Score: 8%**

# Model Tuning

- Unfreeze last few layers
  - When using the pre-trained VGG16 model, all layers are *frozen*, meaning their weights are not updated during training.
  - By unfreezing the last few layers the model is allowed to weights during training.
    - This process helps the model adapt more specifically to the dataset.
- Lower learning rate for fine-tuning
  - This helps to make very small adjustments to the weights of the pre-trained layers.
  - the adjustments to the weights are more subtle, allowing the model to fine-tune its knowledge without forgetting what it has already learned.
- Optimizer: Adam
  - efficient
  - yields good results in a variety of scenarios.



# Misclassifications

True: Common\_Rust  
Pred: corn\_common\_rust



True: Healthy  
Pred: Corn\_\_healthy



True: Gray\_Leaf\_Spot  
Pred: corn\_gray\_leaf\_spot



# Results

Plant Dr: VGG16 Plant Identification and diagnosis model

Unseen Plant Image	Predicted Plant ID and Diagnosis
tomato_late_blight	Strawberry__Leaf_scorch
corn_rust	corn_gray_leaf_spot
apple_scab	Apple__Black_rot

67% Correct Identification and Diagnosis

# Future Plans

- Add more image data
  - Use of an API
  - Image data from more universities/ extension programs
  - Nutrient deficiencies image data
- Try different models
  - Inception (GoogLeNet)
  - MobileNetV2
  - NASNet (Neural Architecture Search Network)
- Creation of an phone app

# References:

- Slide Images:

“Slide 1; Agricultural industry; Farmers”:

<https://www.kdhi-agriculture.com/single-post/agriculture-and-agribusiness-to-drive-industrialization-in-africa>

“Slide 1; Agricultural industry; Nurseries”:

[https://en.wikipedia.org/wiki/Plant\\_nursery](https://en.wikipedia.org/wiki/Plant_nursery)

“Slide 1; Agricultural industry; Research”:

[https://www.seedquest.com/news.php?type=news&id\\_article=120617](https://www.seedquest.com/news.php?type=news&id_article=120617)

“Slide 1: Home Gardening; Retail Stores”: <https://www.facebook.com/FifthSeasonCarrboro/>

“Slide 1: Home Gardening; Hobbyists”:

<https://www.goodhousekeeping.com/home/gardening/g40742429/best-indoor-plants-for-health/>