PlantDx

The Ultimate Mobile Plant Doctor

Jacob Barkow, Kiki Berton, Mima Mirkovic, Payman Roghani, Matt Whittaker
Overview

PlantDx is a modern plant doctor, using **artificial intelligence to diagnose plant health**

Solving two key problems:

1) **Frustration** with the amount of time spent to research and diagnose plant care issues
2) **Cost** to replace dead plants and/or care for dying plants
Problem Space

Opportunity is large and growing, without clear market leader

$25bn +
US indoor plant market size, 2029

- Average household spends approx. $600 on garden supplies per year
- Target market is a subset of plant owners that are digitally-enabled “plant parents”

67%
Adults growing or plan to grow edible plants

- House plant interest surged during pandemic
- We are betting on this trend continuing

47%
People reporting apprehension about purchasing plants due to care concerns ($, time, and uncertainty)

- We strive to solve this problem, freeing up time and saving money

Sources: Data Bridge Market Research Indoor Plants Market, Scotts Miracle-Gro and Wakefield Research, OnePoll
Plant Care: Are You Doing it Right?

7 out of 10 plants die in the care of average plant parents! We understand that finding reliable plant care advice can be challenging.

Due to the number and variety of plant species now available for residential ownership, consumer plant owners often struggle with properly caring for their plants and responding appropriately to the physical signs of potential plant stress.
Dataset / EDA: PlantVillage

Dataset / EDA:
- ~55K w/o augmentation, ~61K w/ augmentation
- ~42K images of healthy and unhealthy labeled plant leaves
- 38 unique classes, ~25 plant types with both healthy and unhealthy labels
- Image size: 256x256

Key Takeaways
- Augmentation requirements for training and generalization (lighting, brightness, etc.)
- Nuanced health characteristics
- Model best suited for image classification is CNN
Modeling approach

- Model selection
  - Focused on Convolutional Neural Networks (CNNs)
  - CNNs have become the standard model for computer vision tasks
- Experimented with 3 different CNN architectures pre-trained on ImageNet data
  - EfficientNet, MobileNet and ShuffleNet
  - Best results achieved with EfficientNet
- Transfer learning: fine-tuned all layers
- Used data augmentation to reduce model overfitting
  - Background transformation
  - The Albumentation library for several image transformations

Final Model: EfficientNet-B0

- Number of parameters: ~5.3M
- Hyperparameters
  - Optimizer: SGD
  - Learning rate: 0.001
  - Batch size: 32
  - Epochs: 10-30
Best performance achieved with moderate data augmentation

Model Training and Evaluation

1. Training on original data
2. Training on moderately augmented data
3. Training on significantly augmented data

Model Performance
Trained on Moderately Augmented Data

Validation Accuracy: 99.384, Validation Loss: 0.019
Model Architecture and Output

1. Plant Species & Health Status
2. Output Confidence (0-100%)
3. Explanation of Diagnosis
4. Actionable Treatment Plan and Recovery Procedure
5. Alternative Diagnoses (not pictured)

We believe this is an unhealthy tomato plant with early blight, with 96.9% confidence. This is a fungal disease that can cause dark spots on leaves and stems. Consider removing infected plant parts and treating with a fungicide.
How did we get here?

Challenges

- Data Availability
  - Solution: Data Augmentation
- Mobile Availability
  - Solution: Mobile App
- Model Deployment
  - Solution: Web App
What’s Next?

Future Work

1. Find deployment solution with greater scalability
2. Optimize model performance
3. DATA!
4. More classes (more plants and diseases)
Mission

Use machine learning to improve plant and consumer happiness outcomes

- Save users time and money
- Increase satisfaction from owning plants

Photo Source: Etsy
Q&A