Plant Disease Prediction Using Machine Learning

Harsha Vardhan Reddy Manam

1)Prototype Selection

a)Problem Statement

The goal is to develop a machine learning-based application that accurately predicts and identifies plant diseases from images. This will enable farmers to diagnose plant diseases early, ensuring timely treatment and potentially saving crop yields.

b)Feasibility, Viability, and Monetization for Plant Disease Prediction System

a. Feasibility:

- Technical Feasibility: The development of a plant disease prediction system using machine learning is feasible within a 2-3 year timeframe. The project requires:
    - Data Collection: Sourcing a robust dataset of plant diseases and symptoms, which can be accomplished through partnerships with agricultural institutions and leveraging existing datasets.
    - Model Development: Developing and training machine learning models using techniques such as convolutional neural networks (CNNs) for image recognition.
    - Application Development: Creating a user-friendly mobile or web application for farmers and agricultural professionals to use.
    - Testing and Validation: Conducting field tests and refining the model based on feedback.
- Resource Availability: Adequate resources, such as skilled data scientists, software developers, and agricultural experts, are available in the current market.
- Funding: Initial funding can be obtained through grants, investments, or partnerships with agribusinesses.

b. Viability:

- Long-Term Relevance: The system is highly relevant for the long term (20-30 years) due to the following reasons:
    - Growing Agricultural Challenges: As climate change and pest pressures increase, the need for efficient disease detection will become even more critical.
    - Technological Advancements: Continuous advancements in machine learning and data analytics will enhance the accuracy and usability of the system.
    - Sustainable Agriculture: The system aligns with global trends towards sustainable farming practices and reducing pesticide usage.
    - Scalability: The system can be scaled to cover a wide range of crops and regions, adapting to new diseases and changing agricultural practices.

c. Monetization:

- Direct Monetization:
  - Subscription Fees: Charging users a monthly or annual subscription fee for access to the app. This can be tiered based on features and user needs (e.g., Basic, Premium, Enterprise plans).
  - Consulting and Expert Services: Offering personalized consulting services and access to agricultural experts for an additional fee.
  - Data Analytics Services: Providing aggregated data analytics to agribusinesses and research institutions.
  - Partnership with Insurance Companies: Offering disease risk assessments and mitigation plans to insurance companies, who can then offer tailored insurance products to farmers.

The direct monetization strategies ensure that the product is not reliant on indirect revenue sources, making it financially sustainable. Each revenue stream is based on providing tangible value to the users and partners, ensuring consistent demand and adoption.

c)Assessment

Market Assessment:

- **Industry Size and Growth:**
  - The global agricultural market is expanding, driven by the increasing demand for food due to population growth.
  - The market for agricultural technology solutions is growing, with a particular focus on tools that improve yield and reduce losses.
- **Regulatory Environment:**
  - Various regions have agricultural regulations emphasizing crop health and disease management.
  - Compliance with agricultural standards and certifications can drive the adoption of advanced disease detection technologies.
- **Competitive Landscape:**
  - While traditional methods of disease detection are common, the integration of AI and ML represents an innovative approach.

    Competing products include manual inspection and traditional diagnostic tools, but this solution offers automated and precise disease identification using image analysis

Customer Assessment:

- **Farmers:**
  - Farmers seek to minimize crop loss and improve yield by detecting diseases early.
- **Agronomists:**
  - Agronomists require accurate and timely disease identification to provide effective crop management advice.
- **Gardening Enthusiasts:**

  - Enthusiasts need easy-to-use tools to maintain the health of their plants.
- **Plant Nurseries:**
  - Nurseries aim to ensure the health of plants being sold and prevent disease spread.

Business Need Assessment:

- **Efficiency and Productivity:**
  - Early disease detection can significantly enhance crop productivity and reduce losses.
- **Cost Savings:**
  - Preventing disease outbreaks can save costs associated with crop replacement, treatment, and loss of yield.
- **Sustainability:**
  - Accurate disease detection supports sustainable farming practices by reducing the need for excessive pesticide use.
- **Market Demand:**
  - There is a growing demand for innovative agricultural technologies that support precision farming and increase overall efficiency.

d) Target Specifications

- **Image Acquisition:**
  - High-resolution image capture of plants for accurate analysis.
- **Machine Learning Models:**
  - Advanced models capable of identifying a wide range of plant diseases from images.
  - Continuous learning and updating of models with new data to improve accuracy.
- **User Interface:**
  - User-friendly interface for easy interaction and interpretation of results.
- **Reporting:**
  - Detailed reports on disease identification and recommended actions for treatment.
- **Integration:**
  - Capability to integrate with existing farm management systems for seamless operation.

e) Customer Characteristics:

- **Farmers:**
  - Looking for reliable and easy-to-use solutions to protect their crops from diseases.
- **Agronomists:**
  - Require detailed and accurate diagnostic tools to support their advisory roles.
- **Gardening Enthusiasts:**
  - Need intuitive and accessible tools for home use.
- **Plant Nurseries:**
  - Focused on maintaining plant health to ensure high-quality sales and prevent disease spread.

f) External Search

Online Information Sources/References:

- https://arxiv.org/abs/2209.07326v3
- https://arxiv.org/abs/1511.08060
  - https://paperswithcode.com/paper/improving-plant-disease-classification-by

2)Prototype Development

Architecture Design:

- Design the system architecture, including the integration of various components such as the user interface, image processing unit, machine learning model, and database.
- Select appropriate cloud services for data storage, processing, and machine learning model deployment.

User Interface Design:

- Develop a user-friendly interface that allows farmers to easily capture and upload images of their plants.
- Design the interface to provide clear and actionable disease identification results, along with treatment recommendations.

Data Processing and Analytics:

- Develop a central data processing unit capable of handling large volumes of image data and performing real-time analysis.
- Ensure the system can handle concurrent users and provide quick response times

Machine Learning Model Development

Data Collection:

- Collect a large dataset of plant images representing various diseases and healthy plants.
- Source data from agricultural research institutions, online databases, and field data collection.

Model Training:

- Train machine learning models using advanced algorithms in deep learning, such as Convolutional Neural Networks (CNNs), to accurately identify plant diseases.
- Implement techniques like data augmentation to enhance the model's robustness and accuracy.

Model Validation:

- Conduct extensive testing and validation of the models using a separate dataset to ensure accuracy and reliability.
- Perform cross-validation to assess the model's performance and prevent overfitting.

System Integration

Component Integration:

- Integrate the image processing unit, machine learning model, and user interface into a cohesive system.
- Ensure seamless communication between the app components and backend services.

User Experience Optimization:

- Develop features for real-time feedback and recommendations based on the model's output.
- Implement notifications and alerts for disease detection and treatment reminders.

Testing and Validation

Simulated Testing:

- Conduct extensive testing in simulated environments to evaluate the system's performance under various conditions.
- Test the app with a diverse set of plant images to ensure robustness and accuracy.

Field Testing:

- Deploy the app in real-world agricultural settings and gather feedback from farmers.
- Validate the app's effectiveness in identifying plant diseases and providing useful recommendations.

Compliance and Standards:

- Ensure the app meets all relevant agricultural and data privacy regulations.
- Conduct security audits to protect user data and ensure compliance with standards.

Key Components of the System :

1. **User Interface (UI):**
    - Mobile app interface for users to capture and upload images of plant leaves.
    - Display area for showing disease identification results and treatment recommendations.
2. **Web Server:**
    - Manages requests between the mobile app and the application server.
    - Ensures secure and efficient data transmission.
3. **Application Server:**
    - Handles image preprocessing and interacts with the machine learning model.
    - Processes the image to prepare it for analysis by the machine learning model.
4. **Machine Learning Model:**
    - Trained to identify various plant diseases from leaf images.
    - Provides predictions based on the input image.

5. **Database:**

    - Stores information about plant diseases, treatment options, and recommended pesticides.
    - Contains details about nearby shops that offer the recommended products.

6. **Backend Services:**

    - Support the application server in preprocessing images and managing data.
    - Ensure the smooth operation of machine learning model inference and data retrieval.

Interaction Flow

1. **Image Upload:**

    - The user captures an image of a plant leaf using the mobile app.
    - The app sends the image to the web server.

2. **Request Handling:**

    - The web server forwards the image to the application server.

3. **Image Preprocessing:**

    - The application server preprocesses the image (e.g., resizing, normalization).

4. **Disease Prediction:**

    - The pre-processed image is fed into the machine learning model.
    - The model predicts the plant disease based on the input image.

5. **Information Retrieval:**

    - The application server searches the database for information related to the predicted disease.
    - Retrieves details about recommended pesticides and nearby shops.

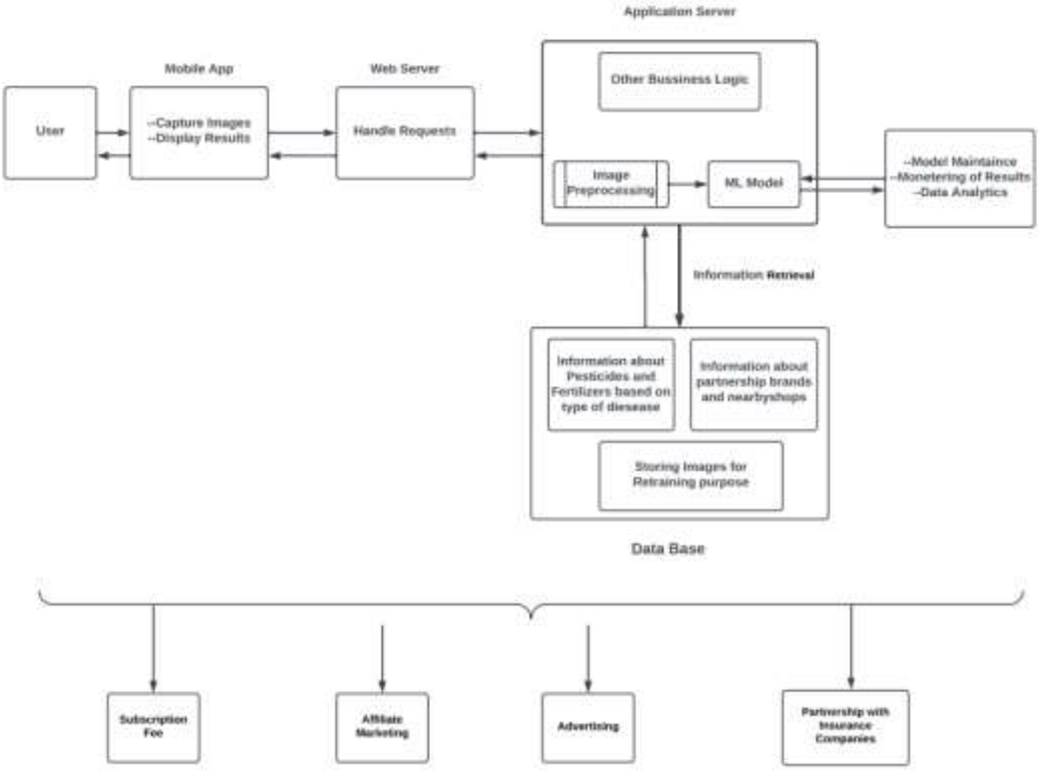
6. **Response Generation:**

    - The application server compiles the prediction results and treatment recommendations.
    - Sends the information back to the web server.

7. **Result Display:**

    - The web server forwards the results to the mobile app.
    - The app displays the disease identification and treatment recommendations to the user.

Schematic Diagram(Prototype Design)



14. Code Implementation

IMPORTING NECESSARY LIBRARIES FOR TRAINING OF MODEL

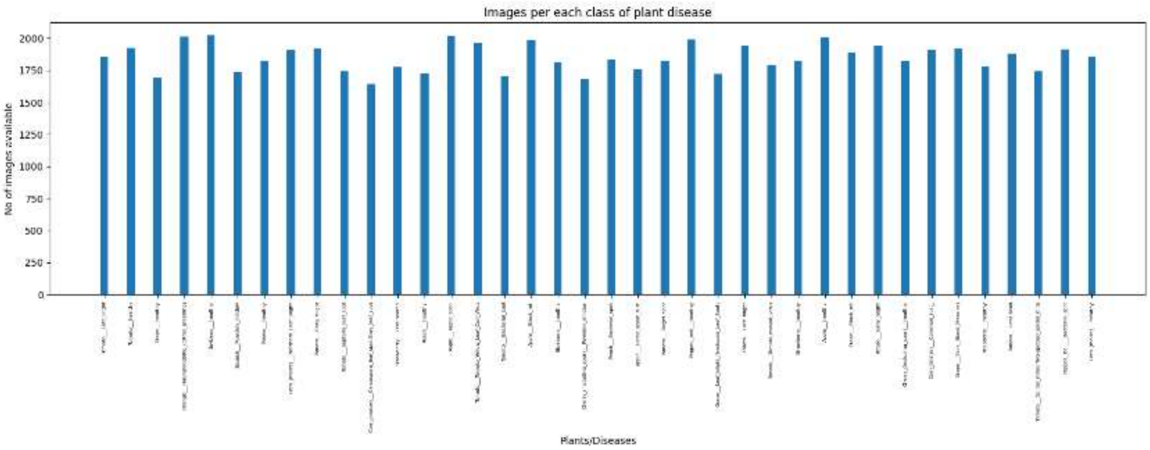
```
import tensorflow
from tensorflow import keras
from keras.models import Sequential,load_model,Model
from keras.layers import Conv2D,MaxPool2D,AveragePooling2D,Dense,Flatten,ZeroPadding2D,BatchNormalization,Activation,Add,In
from keras.optimizers import SGD
from keras.initializers import glorot_uniform
from keras.preprocessing.image import ImageDataGenerator
from keras.callbacks import ModelCheckpoint,EarlyStopping,ReduceLROnPlateau
```

Loading Dataset and Visualizing :

```
path='/kaggle/input/lacaydata/train'
plt.figure(figsize=(70,70))
count=0
plant_names=[]
total_images=0
for i in os.listdir(path):
  count+=1
  plant_names.append(i)
  plt.subplot(7,7,count)

  images_path=os.listdir(path+"/"+i)
  print("Number of images of "+i+":",len(images_path),"||",end=" ")
  total_images+=len(images_path)

  image_show=plt.imread(path+"/"+i+"/"+images_path[0])

  plt.imshow(image_show)
  plt.xlabel(i)

  plt.xticks([])
  plt.yticks([])

print("Total number of images we have",total_images)
```

```python
# plotting number of images available for each disease
index = [n for n in range(38)]
plt.figure(figsize=(20, 5))
plt.bar(index, [n for n in nums.values()], width=0.3)
plt.xlabel('Plants/Diseases', fontsize=10)
plt.ylabel('No of images available', fontsize=10)
plt.xticks(index, diseases, fontsize=5, rotation=90)
plt.title('Images per each class of plant disease')
```

```
Text(0.5, 1.0, 'Images per each class of plant disease')
```



Loading Resnest50 base model and Attaching Custom Classification head :

```python
from tensorflow.keras.applications import DenseNet121
from tensorflow.keras.applications.resnet50 import preprocess_input
```

```python
base_model_tf=DenseNet121(include_top=False,weights='imagenet',input_shape=(224,224,3),classes=num_classes)
```

```python
base_model_tf.trainable = False
pt = Input(shape=(224, 224, 3))
x = base_model_tf(pt, training=False)
x = GlobalAveragePooling2D()(x)
x = Dropout(0.5)(x)
x = Dense(128, activation='relu')(x)
x = Dropout(0.5)(x)
x = Dense(64, activation='relu')(x)
x = Dense(num_classes, activation='softmax')(x)


model_main = Model(inputs=pt, outputs=x)
model_main.summary()
```

Data Augmentation :

```
#Image augmentation
train_datagen= ImageDataGenerator(zoom_range=0.2,width_shift_range=0.2,height_shift_range=0.2,
                                  shear_range=0.2, horizontal_flip=True,fill_mode='nearest

val_datagen=ImageDataGenerator()

path_train='/kaggle/input/lacaydata/train'

path_valid='/kaggle/input/lacaydata/valid'

train= train_datagen.flow_from_directory(directory=path_train,batch_size=32,target_size=(224,224),
                                          color_mode='rgb',class_mode='categorical',seed=42,subset="training", shuffle = Tru

valid=train_datagen.flow_from_directory(directory=path_train,batch_size=32,target_size=(224,224),
                                         color_mode='rgb',class_mode='categorical',subset="validation", shuffle = True)

test= val_datagen.flow_from_directory(directory=path_valid,batch_size=32,target_size=(224,224),
                                       color_mode='rgb',class_mode='categorical', shuffle = False)
```

Model Training :

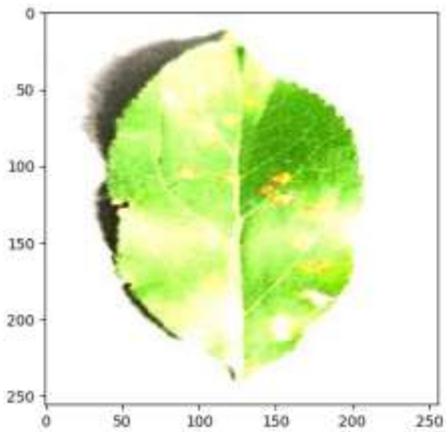
```
model_main.compile(optimizer='Adam',loss='categorical_crossentropy',metrics=['accuracy', 'recall', 'precision', 'f1'])
```

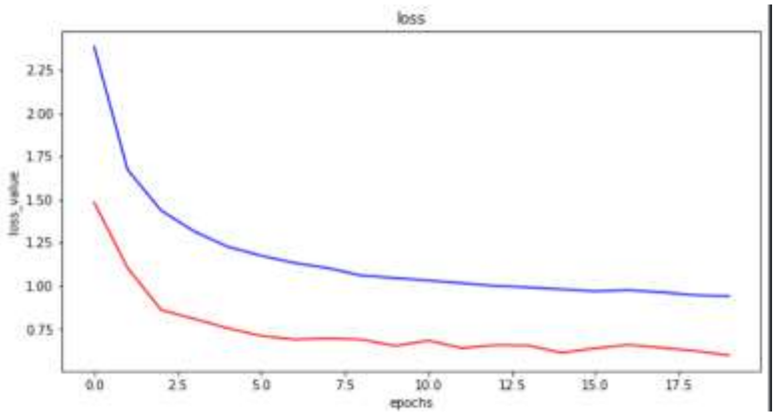
```
history = model_main.fit(train,validation_data=valid,epochs=20,verbose=1,callbacks=[es])
```

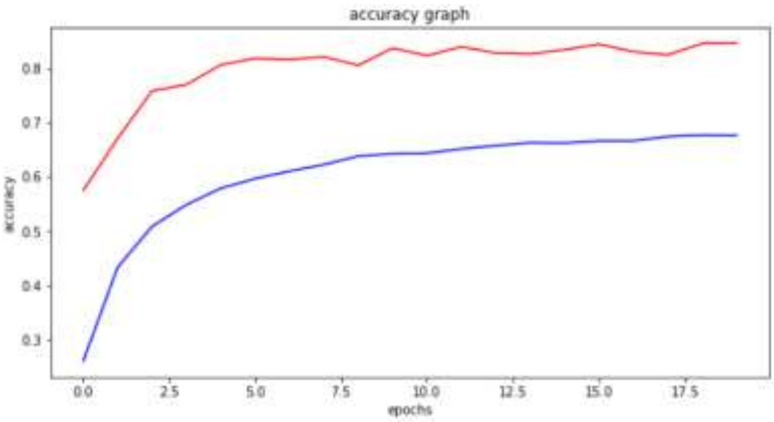
Model Evaluation :

```
# predicting first image
img, label = test[0]
plt.imshow(img.permute(1, 2, 0))
print('Label:', test_images[0], ', Predicted:', predict_image(img, model))
```

```
Label: AppleCedarRust1.JPG , Predicted: Apple___Cedar_apple_rust
```







GitHub Link: https://github.com/HarshaManam49/Plant-Disease-Prediction-/tree/main

3)Business Modelling

1. Value Proposition

- **Early Disease Detection:** Enables farmers to diagnose plant diseases early, ensuring timely treatment and reducing crop losses.
- **Improved Productivity:** Helps in increasing agricultural productivity by maintaining crop health.
- **Cost Savings:** Reduces costs associated with crop loss, replacement, and treatment.
- **Sustainable Farming:** Supports sustainable farming practices by minimizing the use of pesticides.

2. Market Assessment

- **Industry Size and Growth:** The agricultural market is expanding, driven by the need for food due to population growth and the increasing adoption of agricultural technologies.
- **Regulatory Environment:** Compliance with agricultural regulations and standards will drive adoption.
- **Competitive Landscape:** Few direct competitors; traditional methods are prevalent, but this solution offers advanced, automated disease detection.

3. Customer Segments

- **Farmers:** Seeking to minimize crop loss and improve yield.
- **Agronomists:** Require accurate tools for providing effective crop management advice.
- **Gardening Enthusiasts:** Need easy-to-use tools for plant health maintenance.
- **Plant Nurseries:** Aim to ensure the health of plants being sold and prevent disease spread.

4. Revenue Streams

1. **Subscription Fees:**
   - **Basic Plan:** $5/month or $50/year
   - **Premium Plan:** $15/month or $150/year
   - **Enterprise Plan:** Custom pricing
2. **Affiliate Marketing:** Partnering with brands to recommend plant care products, earning a commission on sales.
3. **Advertising:** Integrate agriculture-related ads within the app.
4. **Data Analytics Services:** Offer aggregated and anonymized data services to agribusinesses and research institutions.
5. **Consulting and Expert Services:** Provide consulting services and access to agricultural experts.
6. **Partnership with Insurance Companies:** Offer disease risk assessments and mitigation plans.

7. **Crowdfunding and Grants:** Seek funding from crowdfunding platforms and agricultural grants.

5. Cost Structure

- **Development Costs:** Hiring a development team, purchasing software and hardware, acquiring data.
- **Maintenance Costs:** App updates, server maintenance, customer support.
- **Marketing Costs:** Promoting the app to the target audience.

6. Key Activities

- **Technical Feasibility and System Design:** Assess the technical viability, design system architecture, and user interface.
- **Machine Learning Model Development:** Data collection, model training, validation, and integration.
- **Testing and Validation:** Simulated and field testing to ensure robustness and accuracy.
- **Compliance and Standards:** Ensure the app meets agricultural and data privacy regulations.

7. Key Resources

- **Data Scientists and Machine Learning Engineers:** For model development and training.
- **Software Developers:** With expertise in mobile app development.
- **Agricultural Experts:** For validating disease identification and treatment recommendations.
- **UI/UX Designers:** To ensure the app is user-friendly.

8. Key Partnerships

- **Agricultural Research Institutions:** For data collection and validation.
- **Technology Providers:** For cloud services, data storage, and processing.
- **Agribusinesses and Insurance Companies:** For offering value-added services and risk assessments.

9. Customer Relationships

- **Support and Training:** Provide on-site training, consulting services, and expert consultations.
- **Community Engagement:** Create a user community for sharing insights and feedback.
- **Customer Service:** Offer dedicated account managers for enterprise customers.

10. Channels

- **Mobile Application:** Primary platform for user interaction.
- **Website:** For marketing, support, and additional resources.
- **Partnerships:** Through agribusinesses, insurance companies, and agricultural extension services.

4)Financial Modelling:

Identify which Market your product/service will be launched into

The product/service for plant disease detection will be launched into the following markets:

1. Agriculture and Farming

   - Farmers: Helping them diagnose plant diseases early to prevent crop loss and enhance yield.

   - Agronomists: Providing a tool for precise plant health monitoring and management.

2. Gardening and Horticulture

   - Gardening Enthusiasts: Assisting in the early detection and treatment of plant diseases in home gardens.

   - Plant Nurseries: Ensuring the health of plants before they are sold to customers.

3. Agricultural Technology (AgTech)

   - AgTech Companies: Integrating the machine learning-based system into existing agricultural platforms and solutions.

   - Technology Providers: Offering a diagnostic tool that leverages advanced image processing and machine learning.

4. Research and Education

   - Research Institutions: Providing a tool for studying plant diseases and developing new treatments.

   - Educational Institutions: Aiding in the education of students in plant pathology and agricultural sciences.

5. Government and Non-Profit Organizations

   - Agricultural Extension Services: Supporting government and non-profit initiatives to improve plant health and agricultural productivity.

   - Environmental Organizations: Promoting sustainable farming practices through early disease detection and treatment.

This machine learning-based plant disease detection system aims to address the critical need for innovative, technology-driven solutions in agriculture, contributing to improved crop health, increased yield, and sustainable farming practices.

Collect some data /statistics regarding that Market Online.

Companies under plant disease detection market

Agdia inc:

Agdia has been a leading provider of plant pathogen diagnostics since 1981. Today, the company offers the most comprehensive and trusted portfolio of plant pathogen and GMO testing solutions around the world. Agdia offers a wide range of testing technologies, products, and services aimed at providing clients with the confidence needed to make more educated and informed plant health management decisions. These technologies include, but are not limited to:

- ELISA

- ImmunoStrip

- AmplifyRP isothermal amplification

- Conventional and real-time PCR

- Nucleic Acid Hybridization

- Immunoblot

- ImmunoPrint

- Immunofluorescence

Agdia also has a Testing Services laboratory that can test samples using one or more of the testing methods mentioned above, depending on the pathogen targeted. The company's diagnostic permits allow for the receipt and testing of plants from all over the world.

BIOREBA AG:

BIOREBA AG, established in 1982, is an independent biotech company specializing in diagnostic tests for plant pathogens. Committed to preserving healthy seeds and promoting robust plant growth, BIOREBA's solutions help ensure high crop yields and a stable food supply. The company provides a range of services, including ELISA and PCR tests, AgriStrip rapid diagnostics, and plant sample preparation equipment, with analysis conducted in its ISO/IEC 17025 accredited laboratory.

With operations in over 100 countries, BIOREBA collaborates with research institutions and universities to deliver validated, effective products for various crops. The company's dedicated team and extensive distributor network ensure rapid response and prompt delivery, while its ISO 9001:2015 certification underscores its adherence to rigorous quality standards. This commitment helps farmers worldwide achieve healthy, high-yielding crops.

Creative Diagnostics:
Creative Diagnostics is a leading manufacturer and supplier of antibodies, viral antigens, innovative diagnostic components, and critical assay reagents. The company provides contract biologic R&D and manufacturing services to diagnostic manufacturers, as well as GMP biologics manufacturing for the biopharmaceutical market. Creative Diagnostics aims to be a trusted source for all assay development and manufacturing needs.

The company also supports clients in the rapid development, manufacturing, and commercialization of point-of-care assays, including lateral flow, flow-through, and ELISA tests. Creative Diagnostics specializes in producing fluorescent, visual colloidal gold, and paramagnetic labels, along with highly sensitive, quantitative, and reader-based tests. With decades of experience and unique technologies, Creative Diagnostics is committed to delivering the highest quality products and services to its customers.

Geopard:

GeoPard is an unbiased, cloud-based analytics platform designed for agricultural data. Its core functionality enables the processing of any set of geospatial data, making it a powerful tool for improving agronomic decisions. The platform integrates seamlessly into clients' business processes and applications, with software integration typically completed within a week.

GeoPard offers a comprehensive solution for agricultural operations, covering planning, execution, and practice adjustments based on data insights. It provides cloud analytics for agricultural data, supporting digital farming, agronomy, and precision farming practices. The platform emphasizes accessibility, ensuring that advanced geospatial technological solutions are available at a fair price. By leveraging automation, cloud computing, and open data sources such as satellite imagery, GeoPard optimizes infrastructure costs and delivers modern IT technologies affordably.

Existing Statistics :

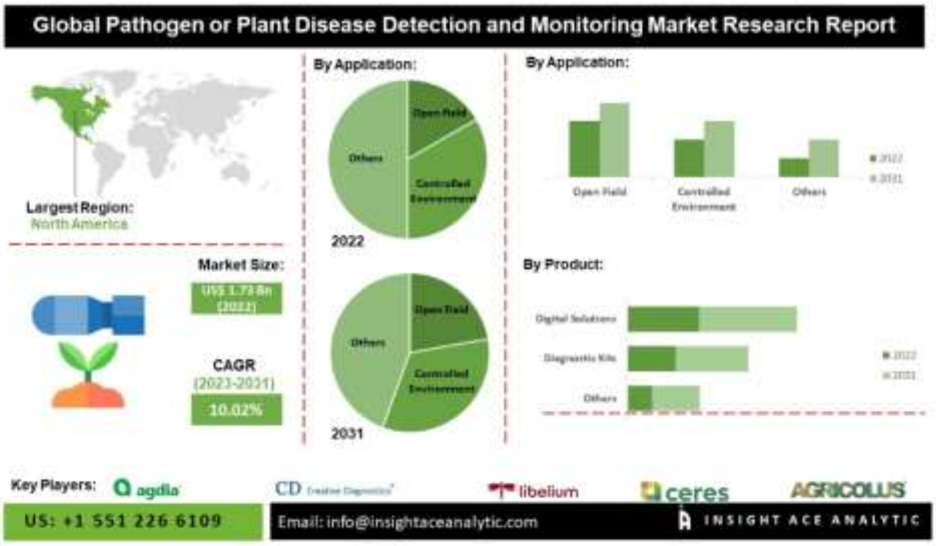
Plant Disease Diagnostics Market Outlook (2023-2033)

Worldwide revenue from the **plant disease diagnostics market** was about **US$ 96.5 Mn** in 2022, with the global market estimated to surge ahead at a CAGR of **4.1%** to reach a valuation of **US$ 150.6 Mn** by the end of 2033.

Attributes	Key Insights
Plant Disease Diagnostics Market Size (2022)	US$ 96.5 Mn
Projected Market Value (2033)	US$ 150.6 Mn
Global Market Growth Rate (2023-2033)	4.1% CAGR
Market Share of Top 5 Countries	51.4%

As assessed by Persistence Market Research, kits are expected to hold a market value of **US$ 43.7 Mn** by 2023. Overall, plant disease diagnostics market sales account for approximately **1.8%** revenue share in the **global agricultural microbials market**, which was valued at around **US$ 5.27 Bn** in 2022.

Pathogen or Plant Disease Detection and Monitoring Market Size is valued at USD 1.73 Bn in 2022 and is predicted to reach USD 3.93 B by the year 2031 at a 10.02% CAGR during the forecast period for 2023-2031.



Financial Equation Corresponding to that Market Trend :

$$\text{Future Value} = \text{Present Value} \times (1 + \text{CAGR})^n$$

where:

- Future Value is the predicted market size in 2031 (USD 3.93 billion).
- Present Value is the market size in 2022 (USD 1.73 billion).
- CAGR is 10.02% or 0.1002.
- n is the number of years (2031 - 2022 = 9).

The financial equation to predict the market size y at any year t can be expressed as:

$$y(t) = \text{PV} \times (1 + \text{CAGR})^{t-2022}$$

Given the values:

- PV = 1.73 billion USD
- CAGR = 0.1002

The equation becomes:

$$y(t) = 1.73 \times (1 + 0.1002)^{t-2022}$$

CAGR :Compound Annual Growth Rate