INTERNATIONAL BACCALAUREATE DIPLOMA PROGRAMME

WORLD STUDIES EXTENDED ESSAY

Discipline: Computer Science and Economics

Area of Study: Science, Technology, and Society

Topic: The Enhancement of Food Security in Ghana through Automated Plant Disease Diagnosis using Machine Learning.

Research Question: How does machine learning model and plant disease diagnosis help to increase food security in Ghana?

Word Count: 3940

TABLE OF CONTENT

Introduction

The global food price index by the U.N Food and Agriculture Organization suggests that an increase in a difficult situation for food security globally especially for low-income consumers who struggle to afford healthy food **(Nordhagen).** Research by Stella Nordhagen - Senior Technical Specialist with the Global Alliance for Improved Nutrition (GAIN) - suggests that the high increase in food prices leading to food insecurity can be caused by the increase in the cost of production such as fertilizers, pests, and pesticides and other factors of production. Food insecurity has been a challenge for a country to feed a growing number of people with a high demand for food. Moreover, this issue can be seen on the national level in Ghana for instance according to the 2012 Ghana Board, "the nation lost $230 million and more than 200,000 tons of cocoa from black pod disease"**("Fighting the Cocoa Black Pod Disease with Drone**

Technology in Ghana"). According to the Ghana Living Standards (2017), "tomatoes make up the highest proportion of 11.6% of household vegetable consumption" **(*GHANA LIVING STANDARDS SURVEY REPORT of the FIFTH ROUND (GLSS 5)*** *Ghana Statistical Service*). Many Ghanaian staple foods include tomato in preparation and as the population is rapidly growing, it is not surprising that the demand for staple crops like tomato will increase. Furthermore, crop production serves as a profitable business and the potential area to increase the livelihood of small-scale farmers to produce large volumes to meet the demand. Unfortunately, farmers face drawbacks in crop production from factors such as late plant disease detection. Plant disease is an alarming issue in the agricultural industry and has a huge impact on crop productivity. Hence if farmers detect the diseases at the early stage, it will help mitigate crop losses therefore resources will be efficiently utilized and no additional cost will be incurred for the treatment of pests and diseases and post-harvest losses. The Ministry of Food and Agriculture (Ghana) goal is to "contribute to sustainable reduction in crop losses caused by pests and diseases which is currently estimated at 30-50%, to about 10-15% with substantially reduced use of hazardous chemicals" **("Plant Protection & Regulatory Services")**. Therefore, the ministry has embarked on educating local farmers on crop pests and disease management. The intervention in conjunction with subsidized fertilizers and improved seeds under the PFJ program increased by 60%, from 18,333MT in 2019 to 29,500MT in 2020. **(Ministry of Food and Agriculture)**

Even Though this solution will have an impact since farmers know pest and disease management, its approaches to disease detection may be time-consuming expensive, and not always accurate. Secondly, the cost spent by the Ministry of Food and Agriculture on pest and disease management could have been used for other activities which is an opportunity cost for instance to fund mechanized tools for farmers to increase their crop yield and productivity. The current manual process of diagnosing plant diseases is not only time-consuming but also prone to errors, and Diagnoses cannot be done on a large scale. Hence using machine learning to develop automated plant diagnoses can help mitigate the problem identified. In this research, I propose that a plant disease diagnosis is developed from convolutional neural networks (CNNs) to identify diseases in real-time. This model relies on images uploaded by farmers and leveraging CNNs helps to learn and extract distinctive attributes from images by employing multiple layers of filters. Subsequently, these attributes from uploaded images are categorized as whether the plant is healthy or not, the type of disease it has, and recommendations to treat the plant. The system is specifically engineered to ensure minimal computational burden, allowing seamless deployment on real-time devices like smartphones which local farmers can use on the field. It is also cost-effective, and time-saving, and precise outcomes render it highly appropriate for utilization.

Furthermore, This model can decrease the environmental and human health issues concerning the widespread use of pesticides. Farmers can reduce the use of insecticides and pesticides by accurately finding the disease at the early stage before the disease increases widespread.

This helped me come up with a research question since machine learning can be a vital solution I can explore to address the issue of food insecurity.

<u>To what extent can machine learning and an automated plant disease diagnosis help to promote food security in Ghana?</u>

Disciplinary Approaches

The research will draw knowledge from two disciplines, computer science and economics. In this case, computer science will enable me to apply technical tools for creating machine learning systems using Python, TensorFlow, and jupyter notebook for image processing and algorithm development. Jupyter Notebook will be a major material for machine learning modeling since it has visualization capabilities and libraries such as Matplotlib, Numpy, and Pandas for data analysis.

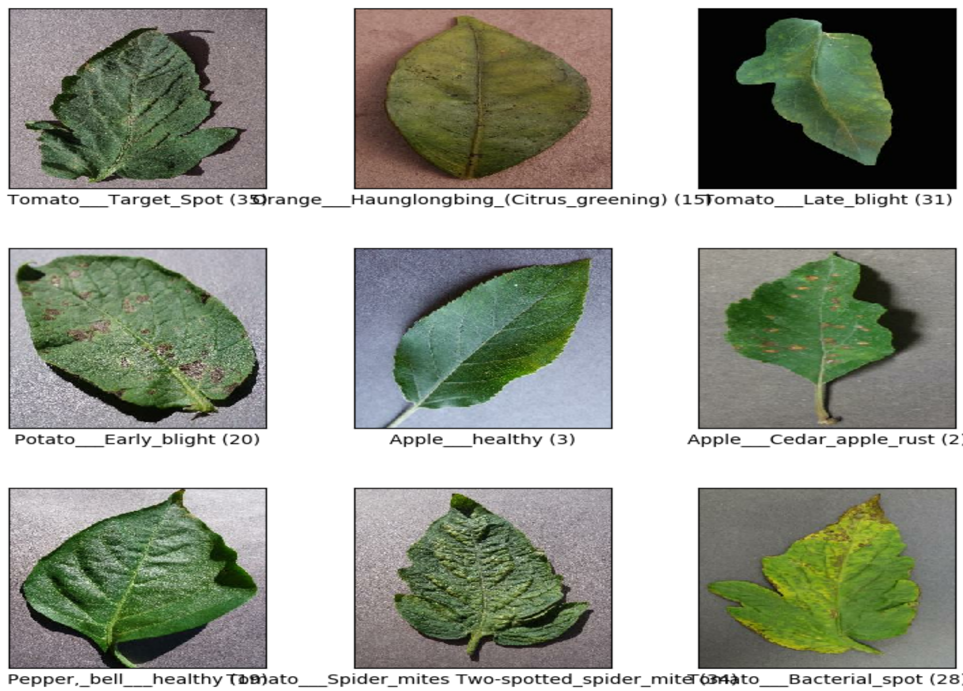
I also aim to use *convolutional neural networks*, a widely used deep learning model, to successfully analyze visual images which can readily be separated into the necessary characteristics. Lastly, a workflow is introduced for our plant disease detection. This field of discipline will enable me to develop a system that farmers can use by uploading real-time images to determine the conditions and health of plants.

Concerning Economics, I will highlight how crop disease diagnoses can efficiently allocate resources such as raw materials, especially pesticides and insecticides. To also evaluate the changes in the supply and demand equilibrium in the market using Aggregate Demand and Supply graphs (AD/AS) and how prices will provide insights into economic impacts. I will be analyzing how farmers respond to incentives for adopting disease detection systems and explain decision-making patterns. Lastly, I will consider how the proposed solution will reduce food insecurity and also promote economic growth using the production possibility frontier in the long run and whether it will promote the increase in the supply of food to meet the increasing demand of consumers since agriculture contributes to 90% of the food needs of the country.

("Ghana at a Glance | FAO in Ghana | Food and Agriculture Organization of the United Nations")

Methodology

To respond to my research question effectively to develop a plant disease detector, I used secondary data which is the plant village data that contains various images of plants affected by diseases and viruses that are common in our crops in Ghana. Therefore, I researched various types of diseases in crops such as tomatoes, maize, oranges, tomato, pepper, and cocoa (https://ipm.ucanr.edu/PMG/diseases/diseaseslist.html). I selected these crops since they are produced in large quantities to provide food for the country. I obtained information from the Ministry of Agriculture on the measures and technology used for pest and disease management. I obtained my dataset of plant images with diseases both healthy and unhealthy for machine learning and deep learning to generate a model for plant disease detection. The dataset was obtained from Tensorflow which consists of 54303 healthy and unhealthy leaf images divided into 38 classes by species and disease **("Plant_village | TensorFlow Datasets").**



(Image Source: Tensorflow)

The plant village dataset has images of 26 different types of plant diseases across crops selected which include tomato, orange, pepper, potato, yam, soybean, and cocoa. Some of the diseases that were included are Black Rot, Bacterial Spot, Late Blight, Leaf Scorch, Early Blight, Powdery Mildew, and swollen shoot virus disease. These diseases are major infections found in Ghana and other West African countries that grow in crops **(Ofori et al.).**

Moreover, I conducted a survey to inquire from farmers to gain insight on factors that affect crop production, technology adoption, and existing solutions implemented and to also gain insight on how it affects the supply of food in the country. The method of investigation will be through questionnaires and surveys and will be responded to by 20 farmers with 10 farmers from the Greater Accra Region, 5 farmers from the Volta region, and 5 farmers from the Ashanti

Region. Secondly, I will use the 2020 report on food security by the Ministry of Finance as secondary data to discuss the impact of inflation caused by an increase in the cost of production and how it affects the affordability of food for local households.

Computer Science

What are Convolutional Neural Networks?

A convolutional neural network is one of the various types of neural networks in deep learning that is tailored to image recognition and processes that involve the manipulation of pixel-level information. They can identify and categorize objects. As a result, it is well-suited for algorithms related to computer vision, where accurately recognizing objects is crucial for plant disease detection.

<u>Diagram of Convolutional Neural Network (figure 2)</u>

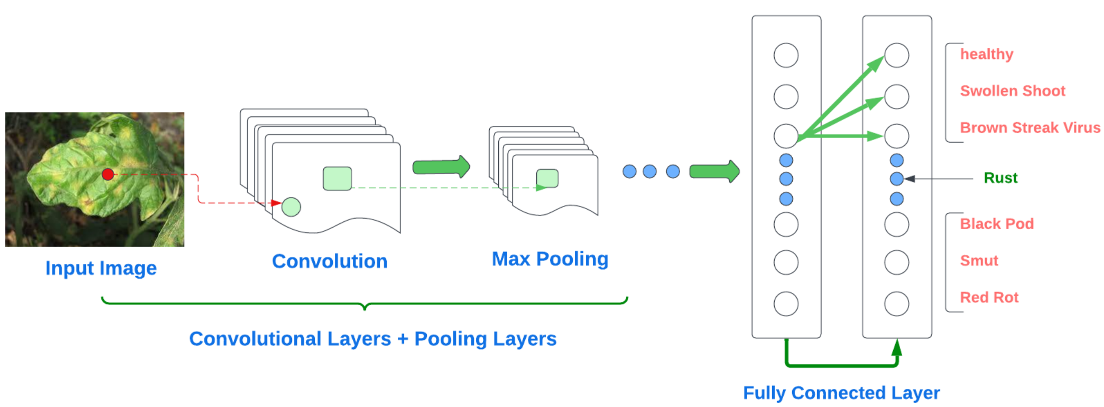
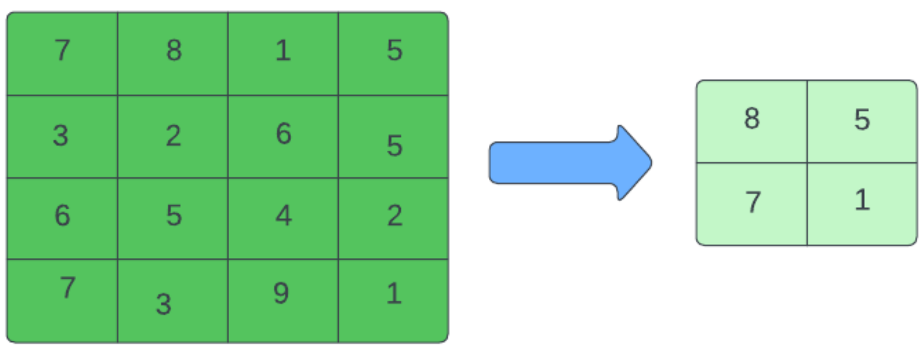


Figure 2 shows the diagram I used in building the convolutional neural network model. The CNN block requires a filter, input data, and a feature map. From the diagram I used a color image labeled as 'Input Image' which is made up of pixels in 3D hence dimensions will have a height, width, and length that relates to RGB. Next, at the convolution process, a filter moves across receptive fields of the image to check if a feature is present. For instance, a filter is applied to an area of the image indicated by a red dot. Then the selected dot area is fed into the output array. Then the filter shifts by a stride and continues until the kernel is swept until the entire image is processed.

Max Pooling Layer

The pooling layer runs over the entire input data then the kernel applies an aggregation function where it looks at the group of data segments and decides on one number to represent them. This way, we get a smaller set of numbers that still give us a good idea of what is in the original data.

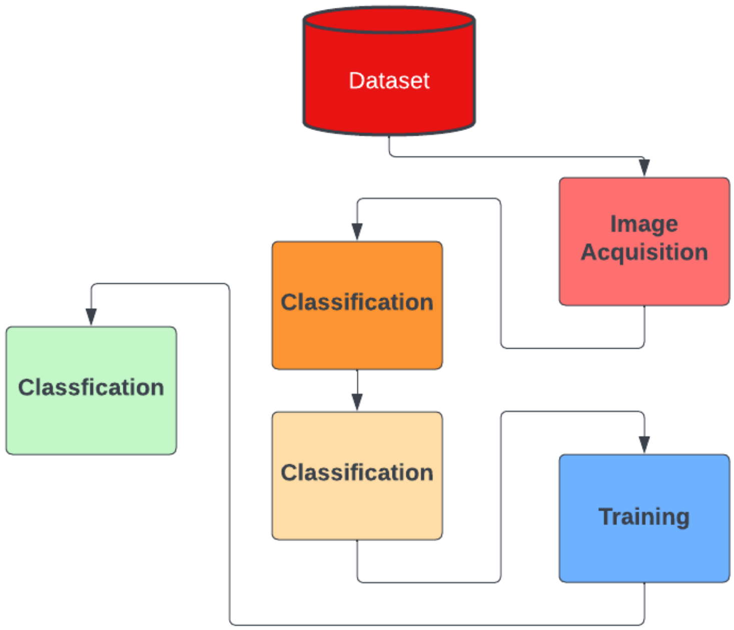


Fully Connected Layer

This layer is the neural network layer and it is always at the end where information is taken from the previous layer and lays it out flat. After, it looks at the flat information and tries to figure out the chances of each possible outcome or class. I used the linear combination and the activation function to create a new output vector thus, an n-dimensional vector, where the N denotes the number of classes of our image classification.

Fig 5: Graph of Number of Images vs Plant Types

The processed dataset of images was trained based on the training model:



From Figure 5, after the image is acquired it goes through the classification process where the feature extraction uses the convolutional and pooling layers. This is where plants are classified whether it is infected with disease or it is healthy.

Preparing Training and Testing Data

I split the data into training and test sets with 80% training and 20% testing, Then training and validation sets were split into 80% training and 20% validation.

Data Preprocessing: Figure 6 shows the preprocessing code for the CNN model

```python
batch_size = 40

train_datagen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=20,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True
)

train_generator = train_datagen.flow_from_dataframe(
    train_df,
    x_col='image_path',
    y_col='label',
    target_size=(224, 224),
    batch_size=batch_size,
    class_mode='categorical'
)

val_datagen = ImageDataGenerator(rescale=1./255)

val_generator = val_datagen.flow_from_dataframe(
    val_df,
    x_col='image_path',
    y_col='label',
    target_size=(224, 224),
    batch_size=batch_size,
    class_mode='categorical'
)

test_datagen = ImageDataGenerator(rescale=1./255)
```

Data preprocessing with a defined batch size

Data augmentation for training set, rescaling pixel values, randomly rotate images b up to 20 degrees

Creating a data generator for the training set, training data frame, column containing image file paths and labels, resize images to pixels

Data augmentation for the validation set

Creating a data generator for the test set

This section of the model defines the data generators for the training, validation, and test sets. These data generators perform data augmentation for the training set and rescaling for all sets. They also specify the target image size and batch size.

Model Selection: After specifying the target image size and batch size, MobileNetV2 architecture is used as the base model with pre-trained ImageNet weights and removes the top fully connected layer.

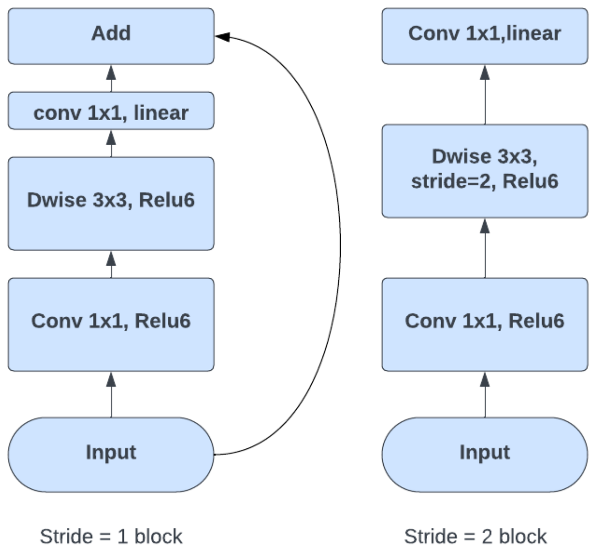


Figure 7: MobileNetV2

With **MobileNetV2,** there are two types of blocks namely stride 1 block and stride 2 block which has the first layer as 1x1 convolution with RELU6. The next layer consists of the depthwise convolution and the last layer is another 1x1 convolution without non-linearity.

Input	Operator	Output
r **x** n **x** t	1 x 1 conv2d, ReLU6	r **x** w **x** (kt)
r **x** n **x** kt	3x3 dwise s=s, ReLU6	$\frac{r}{s}$ **x** $\frac{n}{s}$ **x** (kt)
$\frac{r}{s}$ **x** $\frac{n}{s}$ **x** kt	Linear 1 x 1 conv2d	$\frac{r}{s}$ **x** $\frac{n}{s}$ **x** t'

When there is an expansion factor of k where k = 6 for all experiments If the input is 32 channels, the internal output would be 32 x k = 32 x 6 = 192 channels.

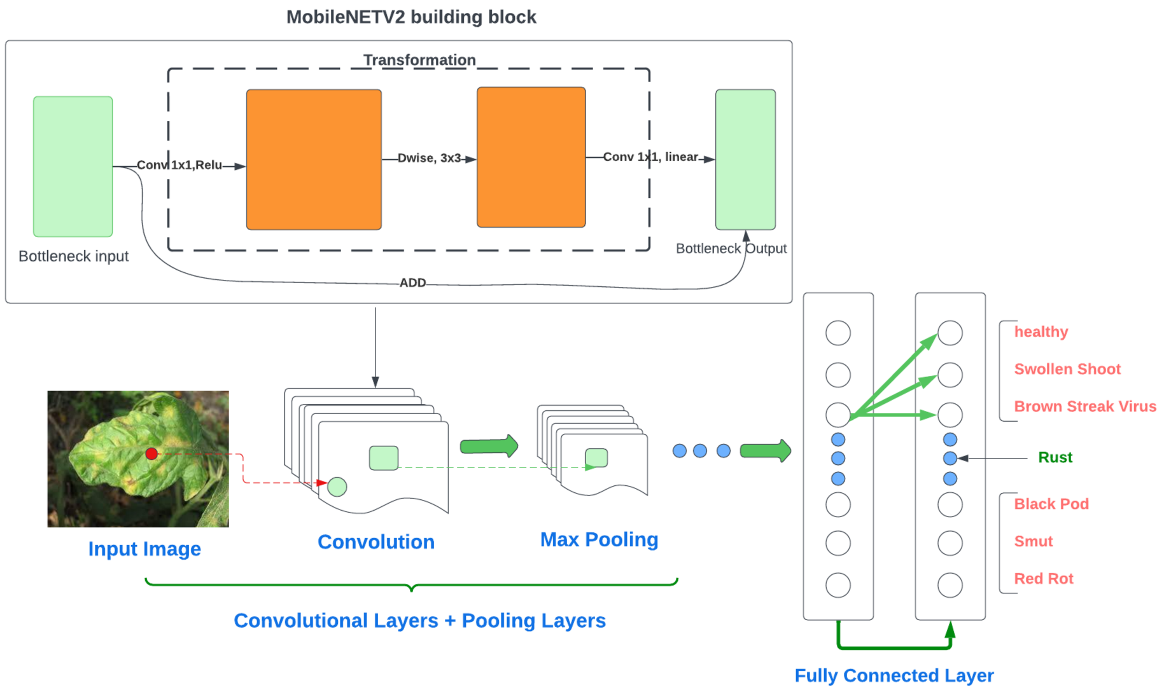


Figure 8: Image showing the MobileNETV2 building block at the convolutional layer

Model Training and Testing:

Figure 9: Image showing the code for model building and training using Adam optimizer, softmax, and epochs

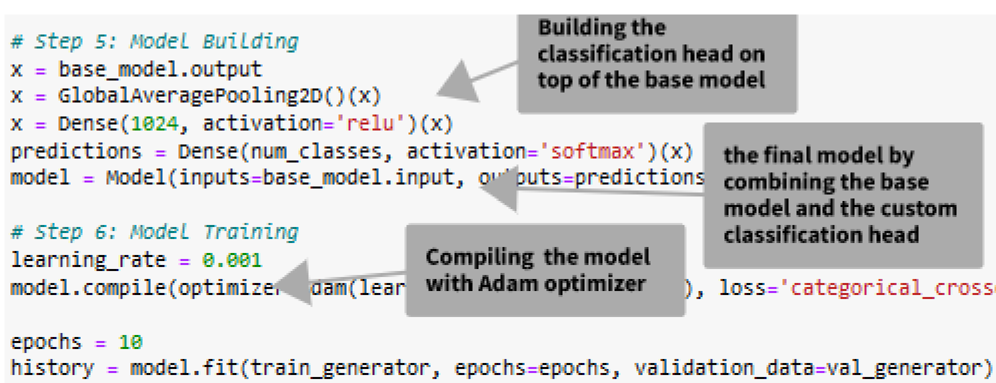
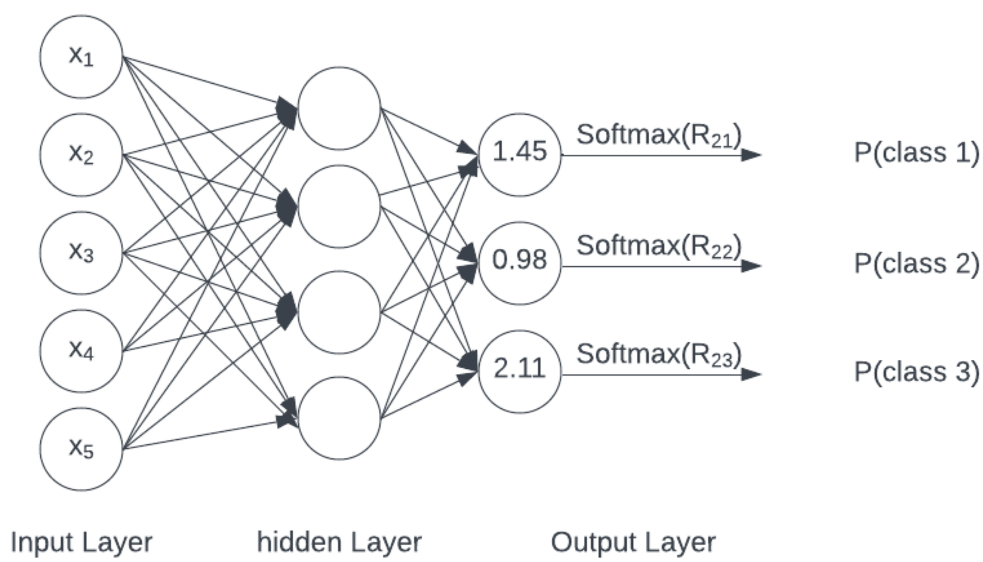


Figure 9 shows how it constructs a custom classification head on top of the base model, consisting of global average pooling, a fully connected layer with ReLU activation, and an

output layer with softmax activation. The softmax activation function uses values of R21, R22, and R23 to determine the final probability value. **("Softmax | What Is Softmax Activation Function | Introduction to Softmax").** Here is the equation for the softmax activation function:

$$softmax(Ri\ ) = \frac{exp(Ri)}{\sum_{j} exp(Rj)}$$

R: The R represents the values from the neurons of the output layer. To normalize and convert them into probabilities, we find the quotient between the R values and the sum of exponential values.



Evaluation Metrics | Limitations

To ensure that the application operates at maximum efficiency, fine-tuning all parameters to achieve accuracy is the most suitable way. Hence I use the loss function called the cross entropy function. This function helps to indicate the amount of uncertainty in machine learning events. The cross-entropy is given by the equation:

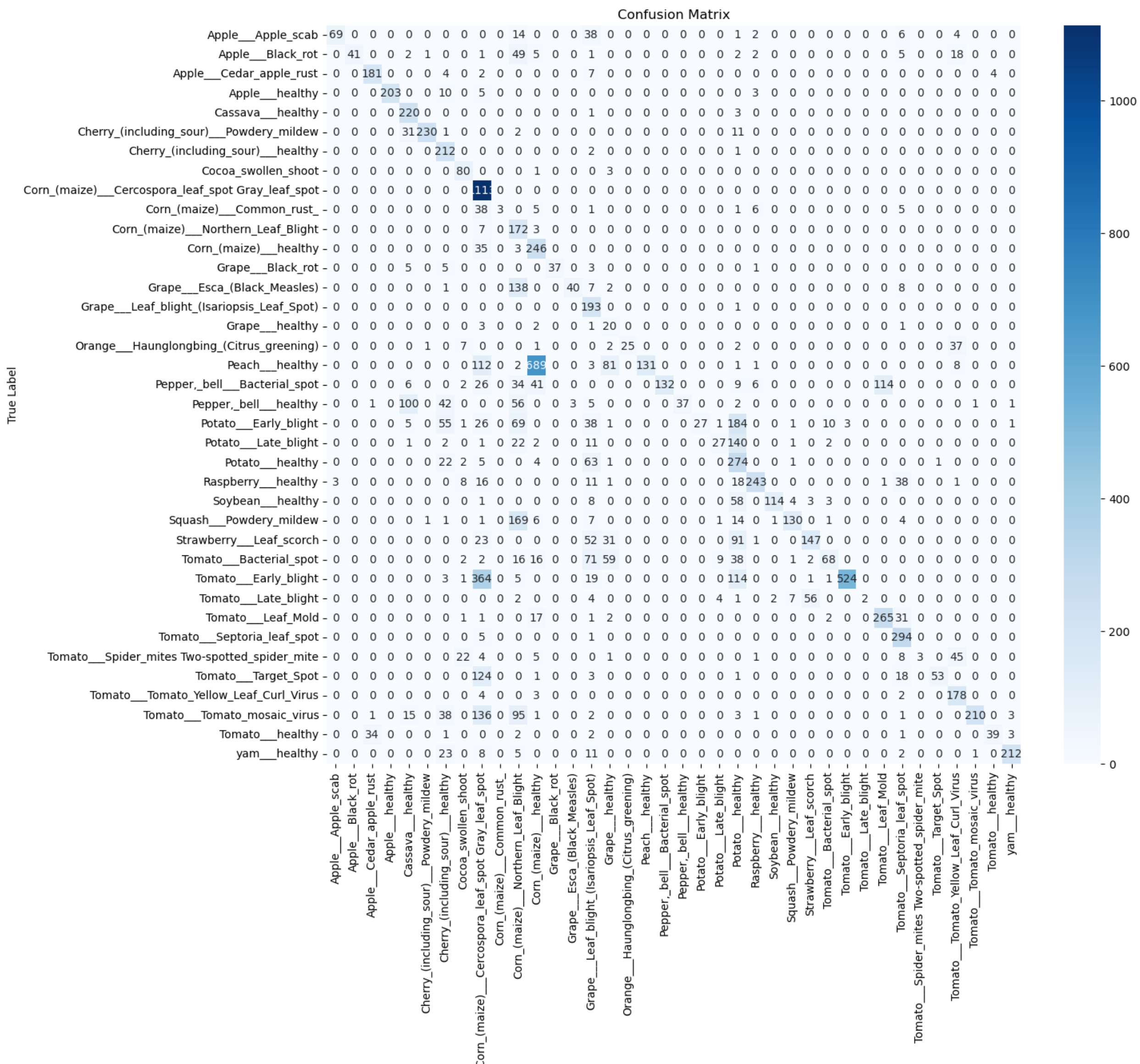
$$H(t,p) \;=\; -\sum_{s\epsilon S} t(s).\,log(p(s))$$

t(s): the true probability distribution (one-shot)

q(x): The model's predicted probability distribution

Confusion Matrix.

Figure 11



A confusion matrix of the CNN was trained on 80% of the images of the dataset and tested on 20% of the remaining images for model accuracy. Figure 11 shows the confusion matrix generated as a result of my machine-learning model. The rows represent the actual classes of image and the columns represent the CNN's class prediction. Each cell in the matrix represents the percentage of images of the row's class that were classified to the column's class. It displays the number of correct predictions: true positives and true negatives and the accuracy of the model can be calculated. When the diagonal of the matrix is higher than its surroundings then it has a higher accuracy and from the diagram, we can see that the highest numbers from Apple_Apple Scab (69) to yam_health (212) create a linear diagonal. Using the Accuracy equation and precision equation of CNN evaluation :

$$\text{Accuracy} = \frac{TRUE\ PREDICTIONS}{ALL\ PREDICTIONS} = \frac{(TP + TN)}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{Actual\ Spam}{Predicted\ Spam} = \frac{TP}{TP+FP}$$

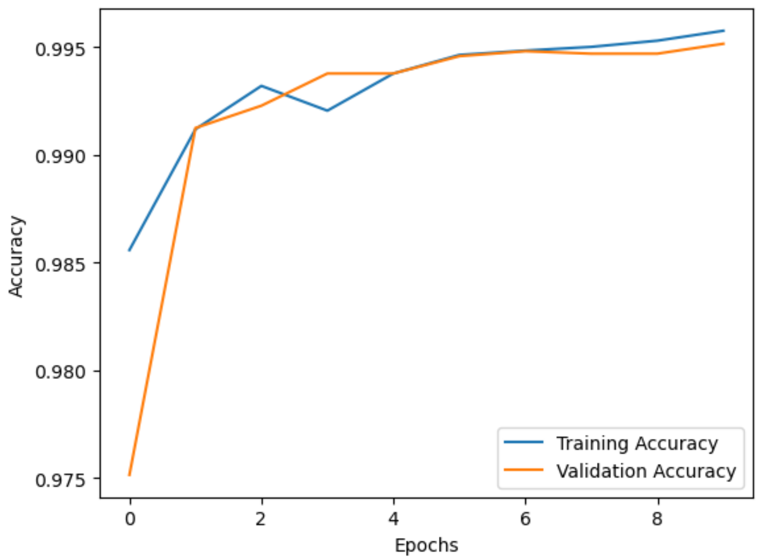
Where: TP = True positive, TN = True negative, FP = False positive, FN = False negative

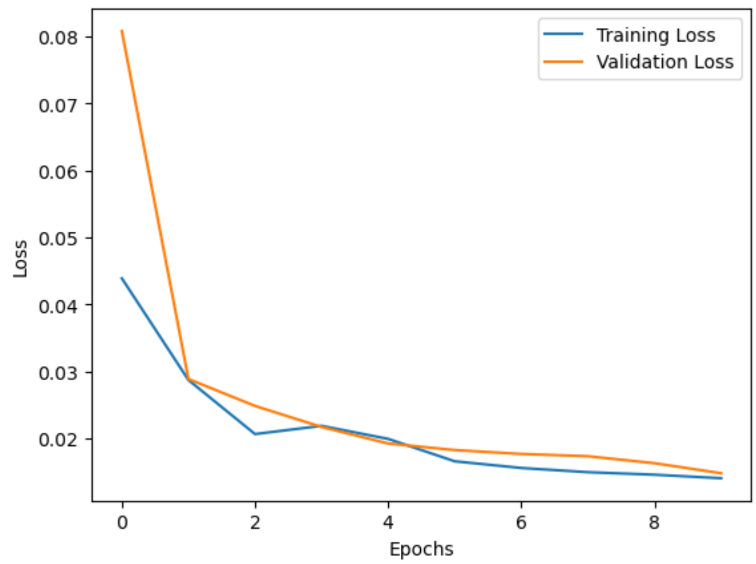
Using Python computation **(Appendix for full code)**, I generated the accuracy of each plant image, and below are the results:

```
# Confusion Matrix and Classification Report
plt.figure(figsize=(14, 12))
conf_matrix = confusion_matrix(test_true_labels, test_predicted_labels)
class_names = list(class_labels_dict.keys())
sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues', xticklabels=class_names, yticklabels=class_names)
plt.xlabel('Predicted Label')
plt.ylabel('True Label')
plt.title('Confusion Matrix')
plt.show();
```

```
print(classification_report(test_true_labels, test_predicted_labels, target_names=class_names))
```

Figure 15





Validation Accuracy Graph **Training Loss and Epochs**

Upon testing the machine learning model with 20% **(refer to the appendix for accuracy result for each plant)** of images from the dataset, Figure 15 shows the training and validation accuracy and training loss to Epochs *(machine learning development time interval).* The nature of the first graph shows a steeply rising curve and plateaus as the epoch increases. The final epoch shows an accuracy close to 0.995 (99.5%) proving the strength of the solution as it has very small wrong predictions. The second graph shows how the model could efficiently learn and improve its understanding with minimal loss as development was going on.

On the other hand, even though it achieved an accuracy of 99.5%, it is still not fully accurate hence farmers will still need some basic knowledge of plant diseases and there may be diseases that may arise making the model outdated hence the model will have to go through retraining and maintenance annually. Secondly, CNN has required a large amount of data and annotation which is time-consuming and costly. They are also prone to overfitting where they store details of training data and fail to generalize to new and different data. Moreover, the plant dataset used for image classification included only some crops such as eggplants and cashew nuts which means may not be used for farmers of these crops.

Overall the model has the potential to solve the problem despite the disadvantages raised from the neural network used.

Economics

What is Food Security?

From the economic point of view, "*food security is defined as when all people at all times, have physical and economic access to sufficient safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life.*" ***(The World Bank)***

Figure 16: Prevalence of Food Insecurity, by Livelihood Group

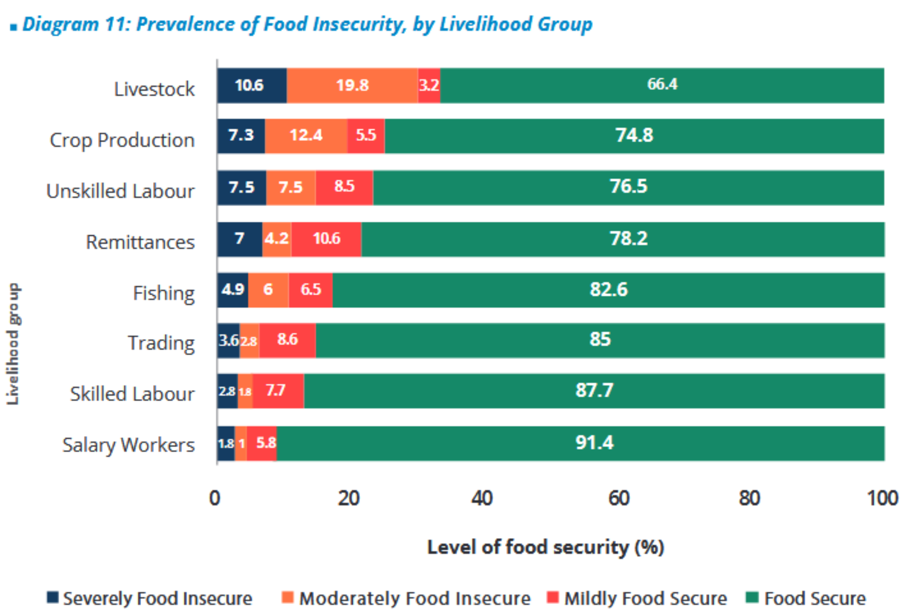


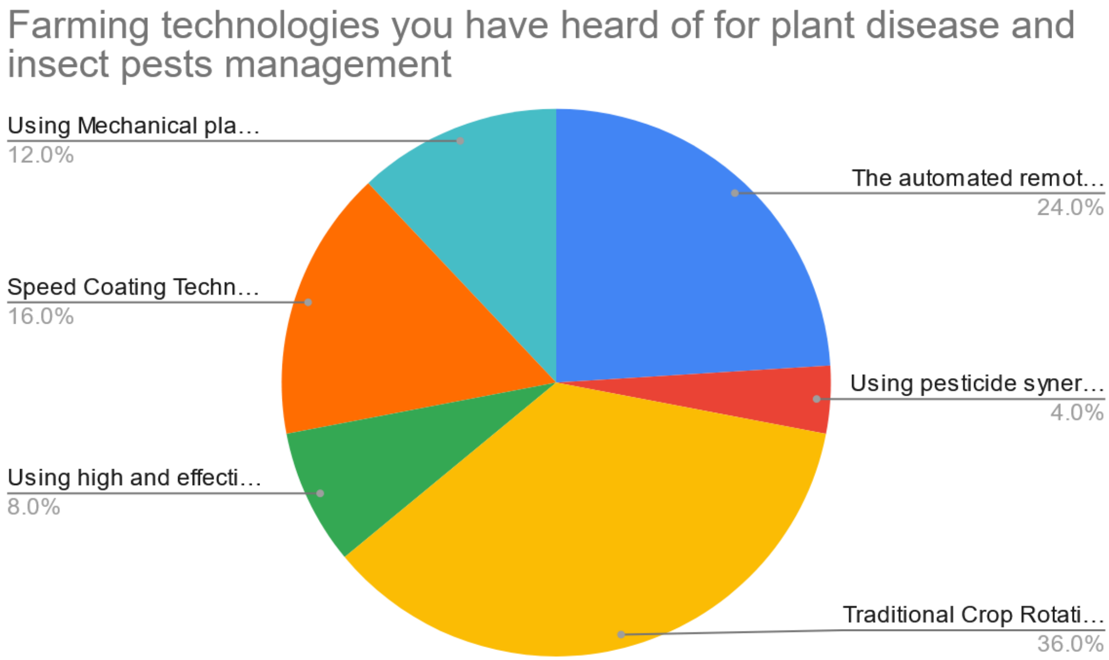
Image Source: ("Ghana - 2020 Comprehensive Food Security and Vulnerability Analysis (CFSVA) - Ghana")

Figure 11 suggests that food insecurity varies greatly among livelihood groups. Firstly Livestock comes with the highest severe food insecurity of 10.6% followed by crop production with 7.3% of severe food insecurity. Hence automated disease detection aids in promoting sufficient safe and nutritious food that meets the preferences of consumers through the reduction of quality and crop losses from plant diseases and pest infestation at an early stage.

Technology Adoption Theory

Technology adoption can be a hindrance to the implementation of automated disease detection in Ghana. For instance, a survey conducted by Abdulai Adams from the Department of Economics at the University of Business and Integrated Development Studies surveyed 461 households of farmers in March 2020 and concluded that 51.4% of the respondents are adopters while 48.6% are non-adopters of the technologies. Therefore, becoming aware of a technology to improve crop production and decrease the cost of production for local farmers is the first step to technology adoption. I surveyed 25 local farmers in the Greater Accra Region, Volta Region, Northern Region, and Ashanti Region to investigate how well they know automated plant disease detection. Below are the results **(Appendix C)**:

Figure 17



From Figure 17, 36.0% of farmers who responded said they had heard of traditional crop rotation to manage plant disease management. Followed by the automated plant disease detection with 24.0%. This shows that farmers are well known about automated plant and pest disease detection however upon further research, 48% of farmers think there is a high risk of implementing plant disease detection for disease and pest control management since there can be inaccuracy and false predictions. However, I believe this model eliminates the assumption of high inaccuracy which is evident from Figure 15 when the training model loss is less than 0.02 and the validation and testing method had a high accuracy of 0.995 to 1.

To increase the adoption of this technology, 33.3% of farmers recommend that farmers should go through retraining to learn the efficiency of implementing it in crop production.

Price Inflation of Food Prices

From the 2020 Comprehensive Food Security and Vulnerability Analysis in Ghana, High food prices after COVID-19 are ranked second among financial shocks that affected the financial stability of 58,588 households. Figure 18 below shows that High food prices caused financial shocks to households showing that households had less disposable income to meet their basic food needs because food prices were high resulted in an increased cost of production of food such as an increase in the cost of fertilizers, pesticides and less supply of food amid the COVID-19 era.

Figure 18

Shock Type	Did not Experience Shocks (%)	Experienced Shocks (%)	Total Number of Households	Ranking by Experience of Shocks
1. COVID-19	36.2	63.8	58,588	1
2. High food prices	65.9	34.1	58,588	2
3. Late rain/drought/no water	78.4	21.6	58,588	3
4. Not enough money to buy food or cover other basic needs	85.7	14.3	58,588	4
5. Reduced income of a household member	86.6	13.4	58,588	5
6. High fuel/transportation prices	88.1	11.9	58,588	6
7. Sudden price fluctuations	88.5	11.5	58,588	7
8. Crop pests/diseases	90.0	10.0	58,588	8
9. Early or heavy rains/floods	93.0	7.0	58,588	9
10. Animal disease/death	95.8	4.2	58,588	10
11. Loss of employment of a household member	96.0	4.0	58,588	11
12. Debt to reimburse	97.2	2.8	58,588	12
13. Delayed pay/salary	97.3	2.7	58,588	13
14. Fire (brush)	98.5	1.5	58,588	14
15. Landslides, (sea) erosion / tidal wave	99.5	0.5	58,588	15

Cost-push Inflation

Cost-push inflation occurs when aggregate supply falls as a result of an increase in the cost of production and in this context, the inflation of food prices is caused by the cost of pesticides, fertilizers, and plant disease treatments.

Figure 19 (author's own)

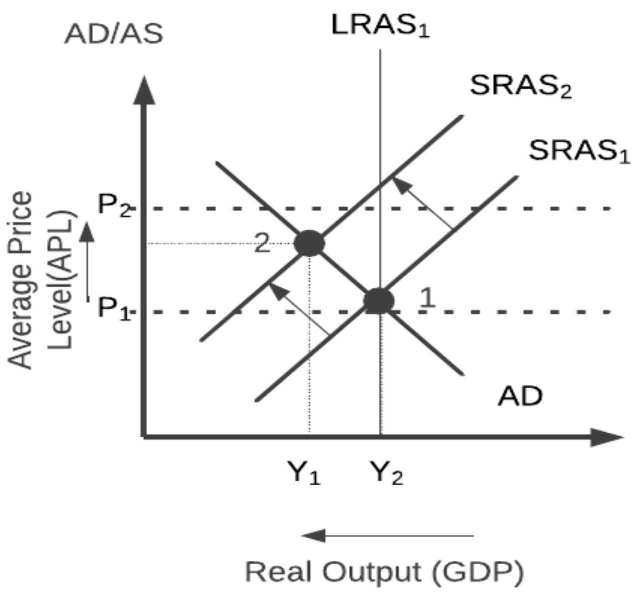
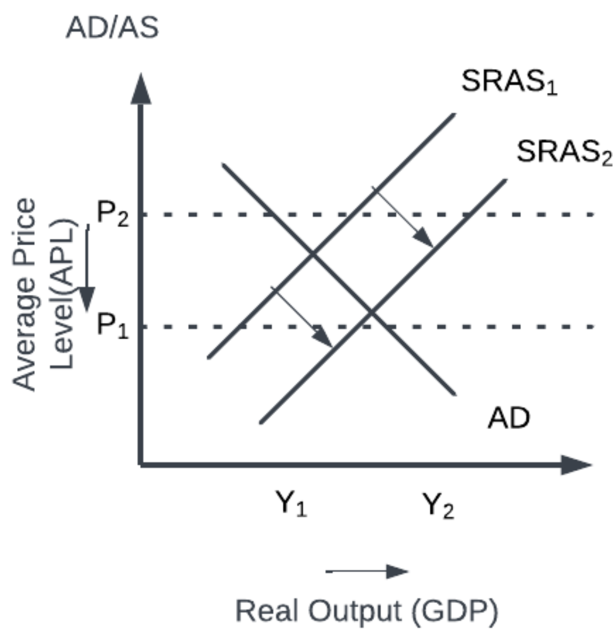


Figure 20(author's own)

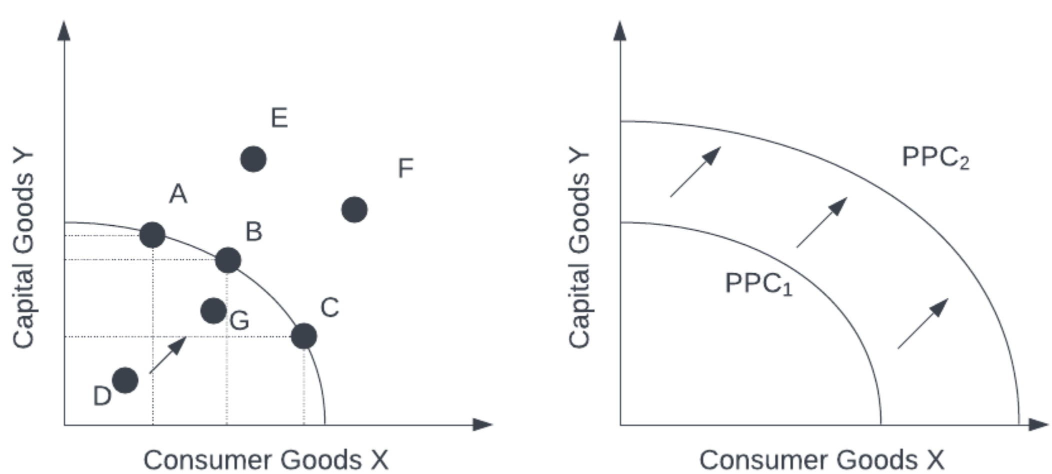


From the diagram above, the increase in the cost of pesticides has caused the shift of the supply curve from $SRAS_1$ to $SRAS_2$ from point 1 to point 2 increasing the price from P_1 to P_2 causing the real output to decrease from Y_2 to Y_1. As the cost of production rises, there may be an increase in unemployment and prices may rise to P_2 to cover these higher costs. A higher price level leading to inflation in the economy caused by the shifts in supply will lead consumers to demand less. Therefore households have less purchasing power to afford food in the basket leading to high food insecurity for each household.

Theoretically, the inflation rate can decrease by shifting the $SRAS_1$ curve to the right which is depicted in **Figure 20** from P_2 to P_1 at a new aggregate supply curve at $SRAS_2$.

The shift of the aggregate supply curve from $SRAS_1$ to $SRAS_2$ caused an increase in real output leading to economic growth. Hence I believe the factor to increase economic growth can be from improving technological methods of farming hence the automated plant disease detection based on a machine learning model can serve as a factor to reduce the cost of production reducing the average price level of food prices since farmers will bear less cost on pest and plant disease control management. The advancement of technology will lead to an increase in local production therefore causing an outward shift in the production possibility curve **(PPC_1 to PPC_2).**

Figure 21: ***Production Possibility Curve (Author's Own)***



At any point on the PPC_1 (A, B, C) shows maximum potential output and actual growth and at points E and F shows potential growth. As a result of the adoption of automated plant disease detection (Technology), it has caused the PPC_1 curve to shift to a new curve at PPC_2 denoting that the economy has increased its actual economic growth.

The government of Ghana has allocated GHC 4,712,927.00 for food security and emergency preparedness and GHC 633,317115.00 for crops including crop production management and livestock development **(*MEDIUM TERM EXPENDITURE FRAMEWORK (MTEF) MINISTRY of FOOD and AGRICULTURE PROGRAMME BASED BUDGET ESTIMATES for 2022*)**. Hence the Ministry of Agriculture with NGO_s intervenes to provide training in IPM, technical backstopping on crop pests and disease identification, and integrated control strategies. Secondly, technical backstopping on the identification of crop presets, training in Good Agricultural Practices (GAP_s), and carrying out diagnostic services. As a result of an increase in government spending to increase food production to promote food security, it has caused a right shift of aggregate demand shown in **Figure 22**. Also, government subsidies, have caused the cost of production to reduce and hence shifted the supply curve to the right from **$SRAS_1$ to $SRAS_2$.** As a result, the GDP will also increase since there are more investments. GDP is based on the following parameters:

GDP = Government Spending + Investment +Consumer Spending + (Export - Import)

Limitation to Methodology

A survey was conducted to find out how farmers use technology in crop production, however, there was regional literacy bias since farmers were not randomly selected from each region but the farmers available.

Conclusions

This essay aims to respond to the question *"To what extent can machine learning and an automated plant disease diagnosis help to promote food security in Ghana?"*. Hence it is important to discuss to what extent various approaches including the government seek to promote food security in Ghana.

Food security is a socio-economic phenomenon in which issues raised should be tackled from various disciplinary approaches. In this report, we discussed how a convolutional neural networks model is used to develop an automated plant disease diagnosis for farmers to be well-known for crop diseases at the early stage. However, this machine learning model is 99.5% accurate which may lead to false predictions. Hence more approaches have to be implemented to tackle the increase in food inflation. Also, the Ministry of Agriculture of Ghana seeks to reduce the cost of production through subsidies and education. With these strategies in place in conjunction with machine learning, food insecurity can be reduced.

Works Cited

"Africa Manages Covid-19; Why Can't It Manage Food Security?" *The Africa Report.com*, www.theafricareport.com/45878/africas-food-security-begins-by-undoing-the-wrongs-of-history/. Accessed 8 Oct. 2023.

"App | Kognity." *App.kognity.com*, app.kognity.com/study/app/2024-economics-hl-2/sid-186-cid-171875/book/the-production-possibility-curve-model-id-29922. Accessed 8 Oct. 2023.

"Cross Entropy Loss: Intro, Applications, Code." *Www.v7labs.com*, www.v7labs.com/blog/cross-entropy-loss-guide.

"Fighting the Cocoa Black Pod Disease with Drone Technology in Ghana." *Mastercard Foundation*, 14 Dec. 2021, mastercardfdn.org/fighting-the-cocoa-black-pod-disease-with-drone-technology-in-ghana/. Accessed 18 Sept. 2023.

"Ghana - 2020 Comprehensive Food Security and Vulnerability Analysis (CFSVA) - Ghana." *ReliefWeb*, reliefweb.int/report/ghana/ghana-2020-comprehensive-food-security-and-vulnerability-analysis-cfsva. Accessed 11 Sept. 2023.

"Ghana at a Glance | FAO in Ghana | Food and Agriculture Organization of the United Nations." *Www.fao.org*, www.fao.org/ghana/fao-in-ghana/ghana-at-a-glance/en/. Accessed 1 Sept. 2023.

GHANA LIVING STANDARDS SURVEY REPORT of the FIFTH ROUND (GLSS 5) Ghana Statistical Service. 2008.

MEDIUM TERM EXPENDITURE FRAMEWORK (MTEF) MINISTRY of FOOD and AGRICULTURE PROGRAMME BASED BUDGET ESTIMATES for 2022.

Ministry of Food and Agricuture. "MEDIUM TERM EXPENDITURE FRAMEWORK (MTEF) for 2021-2024." 2021.

Nordhagen, stella. "Fertilizer Prices Are Rising. Here's What It Means." *Food Tank*, 21 June 2022, foodtank.com/news/2022/06/the-impact-of-higher-input-prices-for-farmers-food-security-and-the-planet/. Accessed 18 Sept. 2023.

Ofori, Atta, et al. "Field Evaluation of the Impact of Cocoa Swollen Shoot Virus Disease Infection on Yield Traits of Different Cocoa (Theobroma Cacao L.) Clones in Ghana." *PLOS ONE*, vol. 17, no. 1, 20 Jan. 2022, p. e0262461, https://doi.org/10.1371/journal.pone.0262461. Accessed 29 Mar. 2022.

"Plant Protection & Regulatory Services." *Mofa.gov.gh*, mofa.gov.gh/site/directorates/technical-directorates/plant-protection-regulatory-services. Accessed 1 Sept. 2023.

"Plant_village | TensorFlow Datasets." *TensorFlow*, www.tensorflow.org/datasets/catalog/plant_village. Accessed 1 Sept. 2023.

"Softmax | What Is Softmax Activation Function | Introduction to Softmax." *Analytics Vidhya*, 5 Apr. 2021, www.analyticsvidhya.com/blog/2021/04/introduction-to-softmax-for-neural-network/. Accessed 8 Oct. 2023.

Tessema, Henok, and Fantahun. *THE EFFECT of POPULATION GROWTH on the ECONOMIC GROWTH of 43 SUB-SAHARAN AFRICAN COUNTRIES: PANEL DATA ANALYSIS (1990-2019)*.

The World Bank. "What Is Food Security?" *World Bank*, 2022, www.worldbank.org/en/topic/agriculture/brief/food-security-update/what-is-food-security. Accessed 8 Oct. 2023.

---. "What Is Food Security? ." *World Bank*, 2022, www.worldbank.org/en/topic/agriculture/brief/food-security-update/what-is-food-security.

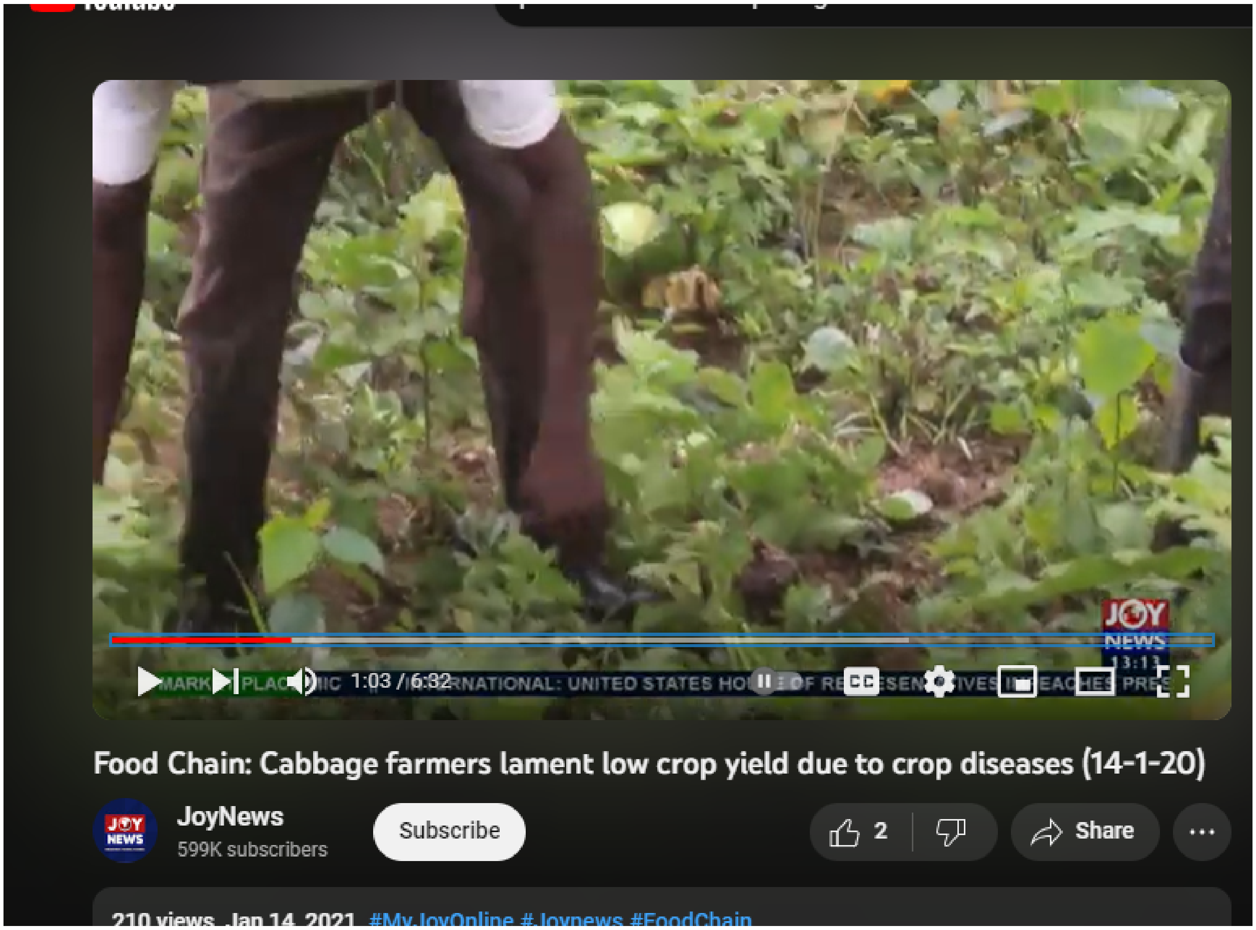
Appendix A

Researcher Reflection Space

Devoting a portion of my daily routine to diligently staying informed about local and global issues around the world brought to my attention how food security in Africa is an alarming concern. Amanda Logan published an article on the Africa repor**t ("Africa Manages COVID-19; Why Can't It Manage Food Security?")** and discusses how Africa is able to manage COVID-19 and still struggles to ensure food security. The worrisome part highlighted was the African stereotype of African limitations stating, "*Africans are portrayed as incapable of producing enough food to feed growing populations*". It was frustrating to see Africa - the land of rich soil- as incompetent to produce enough food for the growing population.

This issue is manifested locally in Ghana. I came across a video by JoyNews (Local News Channel) on YouTube with the headline: "Food Chain: Cabbage farmers lament low crop yield due to crop disease (14-1-20). (https://www.youtube.com/watch?v=rs7hR3RT5mw) . As I watched the video I realized crop disease is an alarming issue in Ghana.

Video source: Youtube (JoyNews)

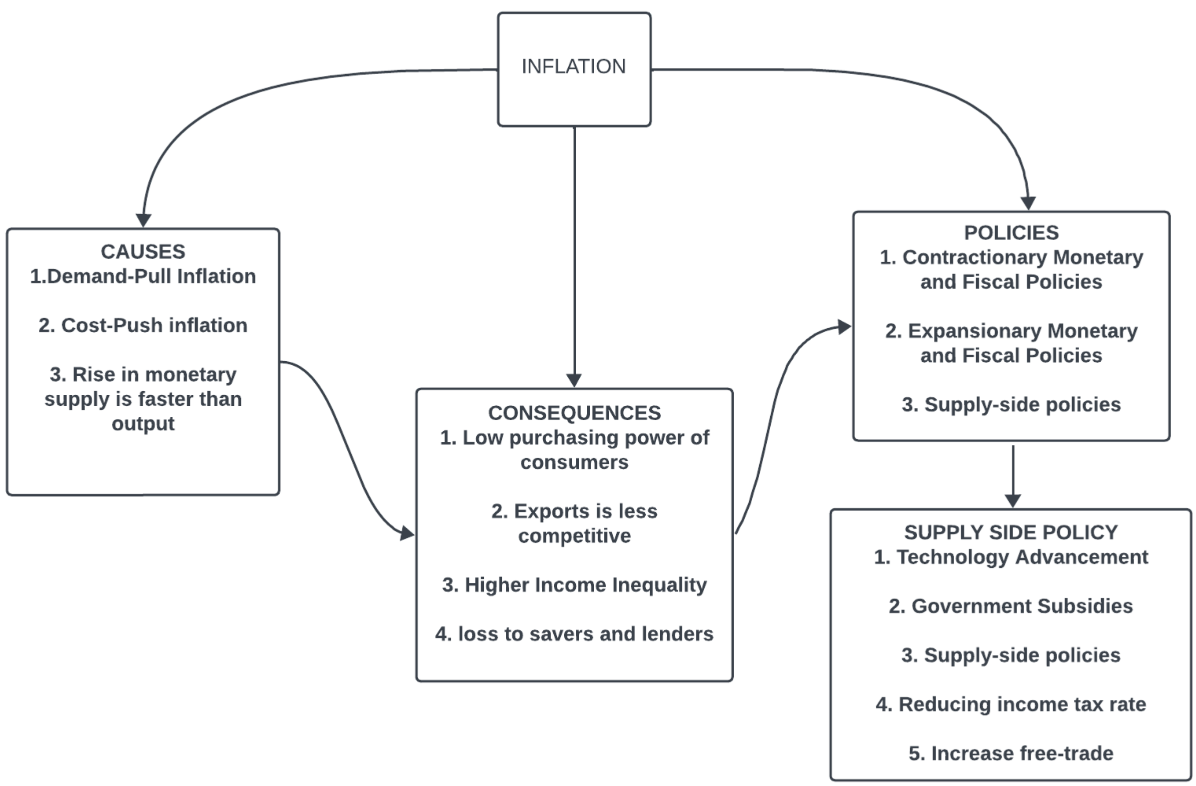


The diamondback moth insect and leave blind disease are destroying farms of local farmers in Armenia central and become freno districts.

A farmer with a half acre of cabbage farm has lost every crop cultivated almost five months ago. They were hopeful that their GHC 5,000.00 investment in this farm would yield enough to cater to their family to generate income. Unfortunately, the crops were infested with leaf blight diseases. The farmers cry to the government to intervene to help them solve this issue. I believe as an IB student who demonstrates the ability of a thinker and is knowledgeable, I can take this as an initiative to help contribute to the solution that can help farmers mitigate the problems of disease infestation on crop yield production.

As an economics student, I learned that since plant disease infestation is a problem that affects the cost of production for farmers, farmers are forced to pass on the cost of production to raise the prices of crops in the market. As a result, an increase in the price of crops can cause an increase in the average price level creating an inflationary gap. Hence it reduces the disposable income of consumers at the low-income level, leading to issues of more consumers not being able to afford food for their households (food insecurity). Furthermore, I learned that to mitigate cost-push inflation, interventions must be implemented on the supply side so I made a mind map of necessary interventions that can be implemented to reduce the cost of production for farmers.

Mind Map (author's own)



Since the cause of the increase in the average price level was from cost-push inflation, I will tackle the problem using the supply-side policy. Out of the policies provided, I am interested in developing a solution using the technological skills I obtained from my computer science lessons, and I will also be pursuing a career in computer engineering. Therefore I brainstormed on effective solutions. I came across an article by the Mastercard Foundation. Three Mastercard Foundation scholars came up with PeculiarAi, an artificial intelligence drone technology that is designed for early detection of the cocoa black pod disease via a mobile application **("Fighting the Cocoa Black Pod Disease with Drone Technology in Ghana")**. However, this system was designed specifically for cocoa and that means other crop production was not eligible for using the application. Hence I decided to tackle inflation and food insecurity through the development of a machine-learning model that can't detect common crops such as cabbage, pepper, tomatoes, and cocoa. This helped me come up with my research question: "How does machine learning plant disease diagnosis help to increase food security in Ghana?".

Appendix B

The Python libraries used in the code you provided are used for the following purposes:

os: interacting with the operating system

numpy: working with n-dimensional arrays

pandas: working with tabular data

matplotlib. pyplot: creating data visualizations

seaborn: creating statistical data visualizations

PIL: working with images

cv2: computer vision

sklearn.model_selection: performing data splitting and hyperparameter optimization

sklearn.metrics: evaluating the performance of machine learning models tensorflow: building and training machine learning model

An Extract from the National Agriculture Investment Plan by Ministry of Agriculture Ghana.

Monitoring and Evaluation Directorate in collaboration with SRID of MoFA will have the responsibility of coordinating the agricultural sector M&E and lead in the collection, collation and analysis of M&E data. The Ministry of Monitoring and Evaluation and the NDPC will play key oversight roles during plan implementation.

A summary of the cost estimates of the plan for the various programmes and sub-programmes are shown below. A total estimated cost of about GH₵9.543 Billion (which excludes that of the cocoa sub-sector strategy) is required for the plan period. It must be noted that the estimated cost of the plan is with respect to public funds. The private sector, especially farmers, will have to invest adequately in the sector to achieve the plan objectives. Development Partners, NGOs and Civil Society Organizations will also have to play their expected roles effectively.

IFJ Budget Estimates (GH₵)

Programme Areas	Indicative Budget (GH₵ 'Million)				
	2018	2019	2020	2021	Total
Programme 1: Management and Administration	**196.30**	**201.33**	**205.98**	**211.10**	**814.72**
SP 1.1: Finance and Administration	154.82	54.82	154.82	154.82	619.28
SP 1.2: Human Resource Development and Management	26.67	29.90	32.89	36.18	125.64
SP 1.3: Policy, Planning, Budgeting, Monitoring and Evaluation	11.85	13.29	14.62	16.08	55,84
SP 1.4: Research, Statistics, ICT and Public Relations	2.96	3.32	3.65	4.02	13.96
Programme 2: Crops and Livestock Development	**747.10**	**2,311.01**	**2,557.34**	**2,800.84**	**7,669.19**
SP 2.1: Crops and Livestock Production and Productivity Improvement	529.57	1,145.95	1,235.08	1,320.30	3,701.34
SP 2.2: Mechanisation, Irrigation and Water Management	160.45	1,049.23	1,229.95	1,423.51	3,702.69
SP 2.3: Post Harvest Management and Agricultural Marketing	54.68	112.94	88.84	52.87	254.65
SP 2.4: Nutrition Sensitive Agriculture	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Programme 2: Crops and Livestock Development	**747.10**	**2,311.01**	**2,557.34**	**2,800.84**	**7,669.19**
SP 2.1: Crops and Livestock Production and Productivity Improvement	529.57	1,145.95	1,235.08	1,320.30	3,701.34
SP 2.2: Mechanisation, Irrigation and Water Management	160.45	1,049.23	1,229.95	1,423.51	3,702.69
SP 2.3: Post Harvest Management and Agricultural Marketing	54.68	112.94	88.84	52.87	254.65
SP 2.4: Nutrition Sensitive Agriculture	2.41	2.89	3.46	4.16	10.51
SP 2.5: Emergency Preparedness	81.64	90.03	99.08	109.04	298.15
Programme 3: Agribusiness Development	**21.41**	**31.20**	**34.59**	**35.95**	**12.15**
SP 3.1: Promotion of Private Sector Investment in Agriculture	0.09	2.39	2.21	1.79	6.48
SP 3.2: Agricultural Finance	21.32	28.81	32.38	34.16	116.67
Programme 4: Sustainable Development of Land and Environment	**201.90**	**222.23**	**244.65**	**267.45**	**936.23**
SP 4.1: Conservation of Natural Resources	4.60	5.06	5.57	6.12	21.35
SP 4.2: Climate Change Mitigation and Resilience Schemes	197.30	217.17	239.08	261.33	914.88
Grand Total	**1,166.71**	**2,765.77**	**3,042.56**	**3,315.35**	**9,543.28**
Government of Ghana Contribution (GOG)	*758.36*	*1,797.75*	*1,977.66*	*2,154.97*	***6,203.13***
Expected Development Partner Contribution (Donor)	*408.35*	*968.02*	*1,064.90*	*1,160.37*	***3,340.15***
Total Secured		***967.84***	***1,150.40***	***1,303.91***	***3,422.16***
GAP (Total not Secured)		**(0.002)**	**(1,892.16)**	**(2,011.43)**	**(6,121.12)**
GoG (MoFA Secured)		471.23	673.66	803.45	**1,948.34**
GAP (GOG)		(1,326.52)	(1,304.00)	(1,351.53)	**(3,982.05)**
Donor (Secured)		496.62	476.74	500.46	**1,473.82**
GAP (Donor)		(471.40)	(588.16)	(659.91)	**(1,719.46)**

- Pesticides and Fertilizer Regulatory
- Plant Quarantine

The Directorate offers Public Services to: Extensionists, Farmers, Importers/Exporters of Agricultural goods, Pesticide Dealers and Applicators, Seed Producers and Dealers, Fertilizer Dealers, Consuming Public and coordinates compliance with obligations in Multilateral Plant Protection Agreements.

Crop Pests Disease Management Division

The Division derives its mandate from Part I of the "Plants and Fertilizer Act" Act 803 (2010)

The Crop Pests & Disease ManagementDivision develops Good Agricultural Practices (GAPs), guidelines for Integrated Pest Management (IPM) of food crops.

The division carries out training and provides comprehensive diagnostic and identification services of plant pests and diseases for stakeholders, monitors the pest situation in the country, ensures effective control of plant pests, manages calamity pest outbreaks (e.g. armyworms, grasshoppers etc), and carries out classical bio-control measures (mass rearing and release of bio-agents), and serves as secretariat for National Fruit Fly Management Committee and National IPM programme.

Services provided:

- **Extension (GO and NGOs)**: Training of Trainers in IPM, technical backstopping on crop pests and disease identification and integrated control strategies.
- **Farmers:** Technical backstopping on pests and disease identification (diagnosis) and integrated control strategies, support in Global GAP certification (training, pre- audit)
- **Exporters of agricultural produce**: Technical backstopping on identification of crop pests, training in Good Agricultural Practices (GAPs) and carries out diagnostic services.
- IGOs: Pests and Disease situation statistics
  Ghana Seed Inspection and Certification Division (GSID)

Appendix C

Below is the full Python code of my machine learning model using the Python and jupyter notebook libraries such as TensorFlow, matplotlib, and Scikeras.

```
# Load the necessary libraries
pip install scikeras
import os
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from PIL import Image
import cv2
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.metrics import classification_report, confusion_matrix,
accuracy_score
import tensorflow as tf
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.applications import MobileNetV2
from tensorflow.keras.layers import Dense, GlobalAveragePooling2D
from tensorflow.keras.models import Model
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.models import save_model, load_model
from tensorflow.keras.preprocessing import image as keras_image
from scikeras.wrappers import KerasClassifier
```

```
data_dir = 'plantvillage dataset\color'

class_folders = os.listdir(data_dir)
image_paths = []
labels = []

for class_folder in class_folders:
    class_path = os.path.join(data_dir, class_folder)
    image_files = os.listdir(class_path)
    for image_file in image_files:
```

```
        image_path = os.path.join(class_path, image_file)
        image_paths.append(image_path)
        labels.append(class_folder)

df = pd.DataFrame({'image_path': image_paths, 'label': labels})
```

```
df.head()
Df.shape
print("The classes:\n", np.unique(df['label']))
# Count the number of images in each class
class_counts = df['label'].value_counts()

# Visualize class distribution using a horizontal bar plot
plt.figure(figsize=(12, 8))
ax = sns.barplot(x=class_counts.values, y=class_counts.index,
orient='h')
plt.title('Class Distribution')
plt.xlabel('Number of Images')
plt.ylabel('Plant Types')
plt.tight_layout()  # Adjust the layout to prevent overlapping labels

# Add data labels to each bar
for i, v in enumerate(class_counts.values):
    ax.text(v + 5, i, str(v), color='black', va='center')

plt.show()
```

```
# Display sample images from each class
num_classes = len(df['label'].unique())
num_images_per_row = 3
num_rows = (num_classes + num_images_per_row - 1) //
num_images_per_row

plt.figure(figsize=(15, 5 * num_rows))  # Adjust figure size based on
the number of rows
```

```
for i, plant_class in enumerate(df['label'].unique()):
    plt.subplot(num_rows, num_images_per_row, i + 1)

    # Inside the loop for displaying sample images
    image_path = os.path.join(df[df['label'] ==
plant_class]['image_path'].iloc[0])

    # Check if the image exists and can be loaded
    if os.path.exists(image_path):
        sample_image = cv2.imread(image_path)
        if sample_image is not None:
            plt.imshow(cv2.cvtColor(sample_image, cv2.COLOR_BGR2RGB))
            plt.title(plant_class)
            plt.axis('off')
        else:
            print(f"Error: Unable to load image from path:
{image_path}")
    else:
        print(f"Error: Image path does not exist: {image_path}")

plt.tight_layout()
plt.show()
```

```
class_labels_dict = {class_label: idx for idx, class_label in
enumerate(np.unique(df['label']))}
df['label'] = df['label'].map(class_labels_dict)
# Split the data into training and test sets (80% training, 20% test)
train_df, test_df = train_test_split(df, test_size=0.2,
random_state=42)
train_df.shape, test_df.shape
# Split the training data into training and validation sets (80%
training, 20% validation)
train_df, val_df = train_test_split(train_df, test_size=0.2,
random_state=42)
train_df.shape, val_df.shape
train_df['label'] = train_df['label'].astype(str)
val_df['label'] = val_df['label'].astype(str)
test_df['label'] = test_df['label'].astype(str)
```

```
print(train_df['label'].unique())


Below is a transcript of the list of questionnaires I used to
interview the local farmers to gain insight into the technology
adoption theory and also the problems of food insecurity in their
lives.
```

RUNNING INITIAL MODEL

```
batch_size = 40

train_datagen = ImageDataGenerator(
    rescale=1./255,
    rotation_range=20,
    width_shift_range=0.2,
    height_shift_range=0.2,
    shear_range=0.2,
    zoom_range=0.2,
    horizontal_flip=True
)

train_generator = train_datagen.flow_from_dataframe(
    train_df,
    x_col='image_path',
    y_col='label',
    target_size=(224, 224),
    batch_size=batch_size,
    class_mode='categorical'
)

val_datagen = ImageDataGenerator(rescale=1./255)

val_generator = val_datagen.flow_from_dataframe(
    val_df,
    x_col='image_path',
    y_col='label',
    target_size=(224, 224),
```

```
    batch_size=batch_size,
    class_mode='categorical'
)

test_datagen = ImageDataGenerator(rescale=1./255)

test_generator = test_datagen.flow_from_dataframe(
    test_df,
    x_col='image_path',
    y_col='label',
    target_size=(224, 224),
    batch_size=batch_size,
    class_mode='categorical',
    shuffle=False
)

# Step 4: Model Selection
base_model = MobileNetV2(weights='imagenet', include_top=False,
input_shape=(224, 224, 3))
# num_classes = len(class_folders)
num_classes = len(class_labels_dict)


# Step 5: Model Building
x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(1024, activation='relu')(x)
predictions = Dense(num_classes, activation='softmax')(x)
model = Model(inputs=base_model.input, outputs=predictions)

# Step 6: Model Training
learning_rate = 0.001
model.compile(optimizer=Adam(learning_rate=learning_rate),
loss='categorical_crossentropy', metrics=['accuracy'])

epochs = 10
history = model.fit(train_generator, epochs=epochs,
validation_data=val_generator)
```

```python
# Check the keys available in the model history
print(model.history.history.keys())
test_loss, test_accuracy = model.evaluate(test_generator)
print(f"Test Loss: {test_loss}, Test Accuracy: {test_accuracy}")
```

ERROR ANALYSIS

```python
# Make predictions on the test set
test_predictions = model.predict(test_generator)
test_predicted_labels = np.argmax(test_predictions, axis=1)

# Get the true labels for the test set
test_true_labels = test_generator.classes

# Create a DataFrame for analysis
error_df = pd.DataFrame({'True Label': test_true_labels, 'Predicted
Label': test_predicted_labels})

# Misclassified images
misclassified_images = error_df[error_df['True Label'] !=
error_df['Predicted Label']]

# Visualization of misclassified images
plt.figure(figsize=(15, 15))
for i, row in enumerate(misclassified_images.head(9).itertuples()):
    img_path = test_df.iloc[row.Index]['image_path']
    img = keras_image.load_img(img_path, target_size=(224, 224))
    plt.subplot(3, 3, i+1)
    plt.imshow(img)
    true_label = class_folders[row._1]  # Use 'True Label' as defined
in error_df
    pred_label = class_folders[row._2]  # Use 'Predicted Label' as
defined in error_df
    plt.title(f'True: {true_label}\nPred: {pred_label}')
    plt.axis('off')
plt.show()

# Confusion Matrix and Classification Report
```

```
plt.figure(figsize=(14, 12))
conf_matrix = confusion_matrix(test_true_labels,
test_predicted_labels)
class_names = list(class_labels_dict.keys())
sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues',
xticklabels=class_names, yticklabels=class_names)
plt.xlabel('Predicted Label')
plt.ylabel('True Label')
plt.title('Confusion Matrix')
plt.show();
print(classification_report(test_true_labels, test_predicted_labels,
target_names=class_names))
```

HYPERPARAMETER TUNING

```
# Unfreeze more layers in the MobileNetV2 model
n = 40
for layer in model.layers[:-n]:
    layer.trainable = True

# Adjust the learning rate for fine-tuning
learning_rate_finetune = 0.00001

# Compile the model with the updated learning rate
model.compile(optimizer=Adam(learning_rate=learning_rate_finetune),
loss='categorical_crossentropy', metrics=['accuracy'])

# Fine-tuning
epochs_finetune = 10
history_finetune = model.fit(train_generator, epochs=epochs_finetune,
validation_data=val_generator)

# Evaluate the Model
test_loss, test_accuracy = model.evaluate(test_generator)
print(f"Test Loss: {test_loss}, Test Accuracy: {test_accuracy}")
# Plot training and validation accuracy
```

```
plt.plot(history_finetune.history['accuracy'], label='Training
Accuracy')
plt.plot(history_finetune.history['val_accuracy'], label='Validation
Accuracy')
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.legend()
plt.show()
# Plot training and validation loss
plt.plot(history_finetune.history['loss'], label='Training Loss')
plt.plot(history_finetune.history['val_loss'], label='Validation
Loss')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()
plt.show()

SAVING THE FINE-TUNED MODEL
# Save the Fine-tuned Model
model.save('fine_tuned_model.h5')

# Load the saved model
loaded_model = load_model('fine_tuned_model.h5')
```

PREDICTION USING THE SAVED MODEL

```
import numpy as np
import matplotlib.pyplot as plt
from keras.preprocessing import image as keras_image
from tensorflow.keras.preprocessing import image as keras_image
# Use the loaded model for making predictions on new images
new_image_path = 'cherry_leaf.JPG'
new_image = keras_image.load_img(new_image_path, target_size=(224,
224))
new_image_array = keras_image.img_to_array(new_image)
new_image_array = np.expand_dims(new_image_array, axis=0)
new_image_array = new_image_array / 255.0  # Normalize the image
```

```python
# Get the predicted class probabilities
predicted_probabilities = loaded_model.predict(new_image_array)[0]
predicted_class_index = np.argmax(predicted_probabilities)

# Define the class_labels_dict with class labels and corresponding
indices
class_labels_dict = {
    0: 'Apple___Apple_scab',
    1: 'Apple___Black_rot',
    2: 'Apple___Cedar_apple_rust',
    3: 'Apple___healthy',
    4: 'Blueberry___healthy',
    5: 'Cherry_(including_sour)___Powdery_mildew',
    6: 'Cherry_(including_sour)___healthy',
    7: 'Corn_(maize)___Cercospora_leaf_spot Gray_leaf_spot',
    8: 'Corn_(maize)___Common_rust_',
    9: 'Corn_(maize)___Northern_Leaf_Blight',
    10: 'Corn_(maize)___healthy',
    11: 'Grape___Black_rot',
    12: 'Grape___Esca_(Black_Measles)',
    13: 'Grape___Leaf_blight_(Isariopsis_Leaf_Spot)',
    14: 'Grape___healthy',
    15: 'Orange___Haunglongbing_(Citrus_greening)',
    16: 'Peach___Bacterial_spot',
    17: 'Peach___healthy',
    18: 'Pepper,_bell___Bacterial_spot',
    19: 'Pepper,_bell___healthy',
    20: 'Potato___Early_blight',
    21: 'Potato___Late_blight',
    22: 'Potato___healthy',
    23: 'Raspberry___healthy',
    24: 'Soybean___healthy',
    25: 'Squash___Powdery_mildew',
    26: 'Strawberry___Leaf_scorch',
    27: 'Strawberry___healthy',
    28: 'Tomato___Bacterial_spot',
    29: 'Tomato___Early_blight',
    30: 'Tomato___Late_blight',
    31: 'Tomato___Leaf_Mold',
```

```python
    32: 'Tomato___Septoria_leaf_spot',
    33: 'Tomato___Spider_mites Two-spotted_spider_mite',
    34: 'Tomato___Target_Spot',
    35: 'Tomato___Tomato_Yellow_Leaf_Curl_Virus',
    36: 'Tomato___Tomato_mosaic_virus',
    37: 'Tomato___healthy'
}

# Check if the predicted_class_index is present in the dictionary
if predicted_class_index in class_labels_dict:
    predicted_class_label = class_labels_dict[predicted_class_index]
else:
    predicted_class_label = 'Unknown Class'

# Display the image and prediction
plt.imshow(new_image)
plt.title(f"Predicted Class: {predicted_class_label}")
plt.axis('off')
plt.show()
```

ACCURACY RESULT (precision, recall, f1-score and support)

	precision	recall	f1-score	support
Apple___Apple_scab	0.96	0.51	0.67	134
Apple___Black_rot	1.00	0.32	0.49	127
Apple___Cedar_apple_rust	0.83	0.91	0.87	198
Apple___healthy	1.00	0.92	0.96	221
Cassava___healthy	0.57	0.98	0.72	224
Cherry_(including_sour)___Powdery_mildew	0.99	0.84	0.91	275
Cherry_(including_sour)___healthy	0.50	0.99	0.67	215
Cocoa_swollen_shoot	0.63	0.95	0.76	84
Corn_(maize)___Cercospora_leaf_spot Gray_leaf_spot	0.54	1.00	0.70	1113
Corn_(maize)___Common_rust_	1.00	0.05	0.10	59
Corn_(maize)___Northern_Leaf_Blight	0.20	0.95	0.33	182
Corn_(maize)___healthy	0.23	0.87	0.37	284
Grape___Black_rot	1.00	0.73	0.84	51
Grape___Esca_(Black_Measles)	0.93	0.20	0.33	196
Grape___Leaf_blight_(Isariopsis_Leaf_Spot)	0.34	0.99	0.51	194
Grape___healthy	0.10	0.74	0.17	27
Orange___Haunglongbing_(Citrus_greening)	1.00	0.33	0.50	75
Peach___healthy	1.00	0.13	0.23	1028
Pepper,_bell___Bacterial_spot	1.00	0.36	0.53	370
Pepper,_bell___healthy	1.00	0.15	0.26	248
Potato___Early_blight	1.00	0.06	0.12	422
Potato___Late_blight	0.64	0.13	0.22	209
Potato___healthy	0.28	0.73	0.41	373
Raspberry___healthy	0.91	0.71	0.80	340
Soybean___healthy	0.97	0.60	0.74	191
Squash___Powdery_mildew	0.90	0.39	0.54	336
Strawberry___Leaf_scorch	0.70	0.43	0.53	345
Tomato___Bacterial_spot	0.78	0.24	0.37	284
Tomato___Early_blight	0.99	0.51	0.67	1032
Tomato___Late_blight	1.00	0.03	0.05	78
Tomato___Leaf_Mold	0.70	0.83	0.76	320
Tomato___Septoria_leaf_spot	0.69	0.98	0.81	300
Tomato___Spider_mites Two-spotted_spider_mite	1.00	0.03	0.07	89
Tomato___Target_Spot	0.98	0.27	0.42	200
Tomato___Tomato_Yellow_Leaf_Curl_Virus	0.61	0.95	0.74	187
Tomato___Tomato_mosaic_virus	0.99	0.42	0.58	506
Tomato___healthy	0.91	0.48	0.62	82
Yam___healthy	0.96	0.81	0.88	262

This is the transcript of the survey I created to get responses from local farmers in Ghana in other to gain information about technology adoption theory and the implication of cost of production on the income of households and its effect of food security.

Supplementary Materials:

Farmers' Machine Learning Plant Disease Detection Questionnaire

I am very grateful for your contribution in my academic research

____

Name of respondent: _____________________________

Region:_______ Town: _______ Village: ______ Ethnic Group: ________

____

Basic information about the farmer's household

Total household size:________________

Family Member Code : 1 2 3

Relationship with head of household

A = Household head; B = Couple; C = Father and son; D = Household Mother and son; E = Father and daughter; F = Mother and head(A) daughter; G = Other

Respond: ____________

Gender

A = Male; B = Female

Answer: ___________

Age

A = 20 and below; B = 21-30; C=31-40; D=41-50; E=51-65; F = 66 above

Answer:

Education level

A = No or little literacy; B = Primary education; C = Junior/middle school; D = Technical secondary and high school; E = College and above

Answer: ____________

Years of crop production

A = 10 years and below; B = 11-20; C = 21-30; D = 31-40; E = 41-50; F = 51 years and above

Answer: ____________

II Scale of family business and income and expenditure (2023)

1. How is the contracted land handled now?
    a. Own business
    b. Abandoned
    c. Free transfer to others to plant
    d. Leased to others for a fee____ for ___ years/months
    e. Other

2. Your farm's largest agricultural expenditure is GHC ________. The largest agricultural expenses are mainly for ()
    a. Fertilizer
    b. Pesticides
    c. Seed Purchase
    d. Agricultural machinery for cultivation and harvesting
    e. Other

3. The main type of agricultural business of your household is ()
    a. Crops : Cocoa, pepper, tomato, oranges and others
    b. Livestock and poultry
    c. Other

III. Adoption and Perception of Machine Learning Automated Plant Disease Detection Technology (2023)

1. Please tick the options of farming technologies you have heard of for plant disease and insect pest management
    A. Integrated Control of Plant Protection
    B. Traditional Crop Rotation Method
    C. Speed Coating Technology
    D. The automated remote plant disease detection monitoring for plant diseases and insect pest infections.
    E. Using high and effective pesticides such as xiaochongthion, Chlorpyrifos, and butylene fipronil.
    F. Using pesticide synergists and growth regulators such as compound sodium nitrophenolate, brassins, and gibberellic acid.
    G. Using Mechanical plant equipment to apply the right volume of pesticides precisely

2. On a scale of 1-10 how well do you know about the automated plant disease diagnosis using a trained machine learning model?

1 2 3 4 5 6 7 8 9 10

3. Do you need such technologies?
    A. Yes
    B. No

4. Why do you need them?
    A. Protect the environment
    B. Increase income
    C. Produce healthy crops with less toxicity from pesticides
    D. Others are using it
    E. Improve Yield and Quality

5. Which medium do you think is best to learn about the technology?

A. Training on how to implement the monitoring plant disease detection system
B. Oral recommendations from friends and neighbors
C. Learning from farmers using it in their production
D. Lectures by relevant experts from government institutions
E. News on radio, TV, newspapers, and other social media
F. Books and technical manuals

6. Do you use this technology?

A. Yes
B. No

7. What technology would you like to implement on your farm?

A. Integrated Control of Plant Protection
B. Traditional Crop Rotation Method
C. Speed Coating Technology
D. The automated remote plant disease detection monitoring for plant diseases and insect pest infections.
E. Using high and effective pesticides such as xiaochongthion, Chlorpyrifos, and butylene fipronil.
F. Using pesticide synergists and growth regulators such as compound sodium nitrophenolate, brassins, and gibberellic acid.
G. Using Mechanical plant equipment to apply the right volume of pesticides precisely

8. Are you satisfied with the programs implemented by the Ministry of Food and Agriculture in helping increase crop yield production to promote Food Security?

A. Yes
B. No

9. Do you think the impact of adopting automated plant disease detection on crops is safe?

   A. Very safe
   B. Safe
   C. Neutral
   D. Unsafe
   E. Very Unsafe

10. Do you think the risk of adopting automated plant disease detection on crops is large?

    A. Yes
    B. No
    C. Maybe

11. What are some government policies you know about such as government subsidies for pesticides, training programs for plant disease and pest control management, and subsidies for chemical fertilizer technology? Are you satisfied? If not, why?

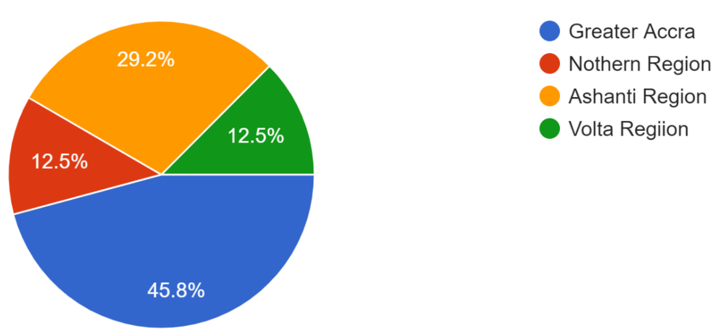
    A. Very satisfied
    B. Satisfied
    C. Neutral
    D. Dissatisfied
    E. Very dissatisfied

Appendix D

The Summarised data of the questionnaire result:

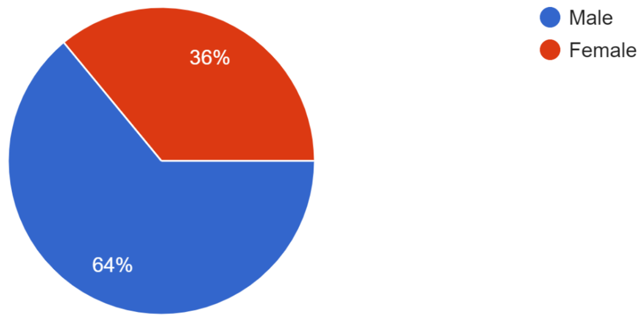
Region

24 responses



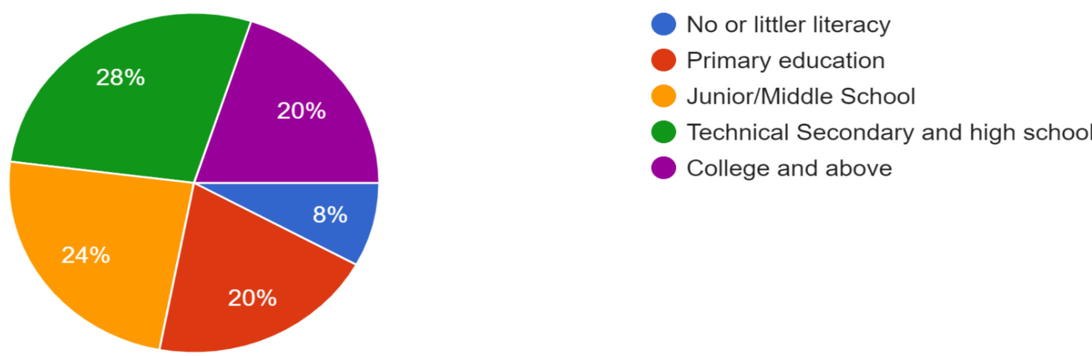
Gender

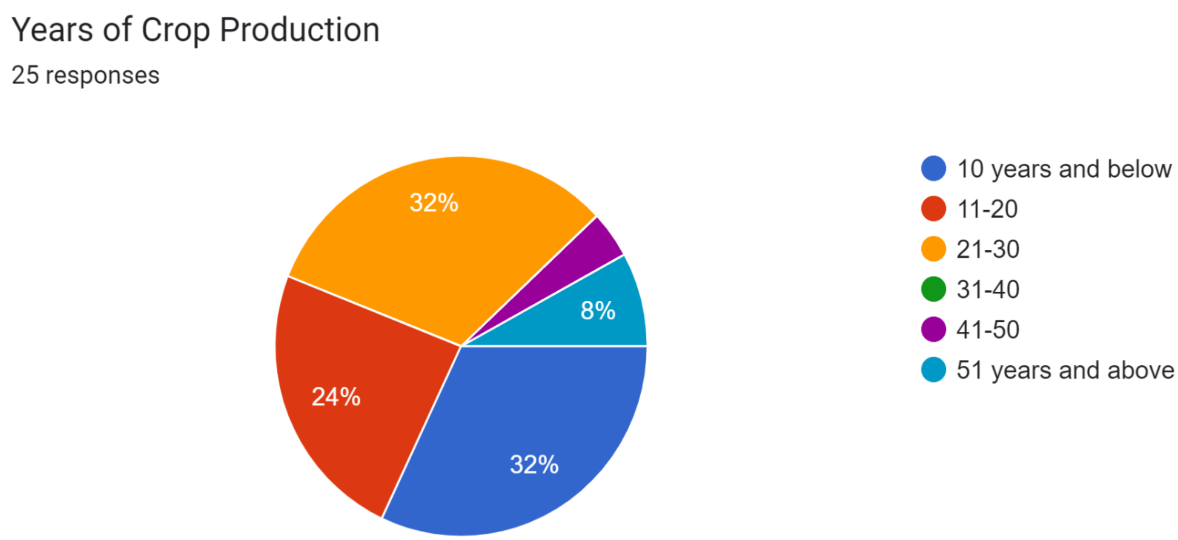
25 responses



Education Level

25 responses

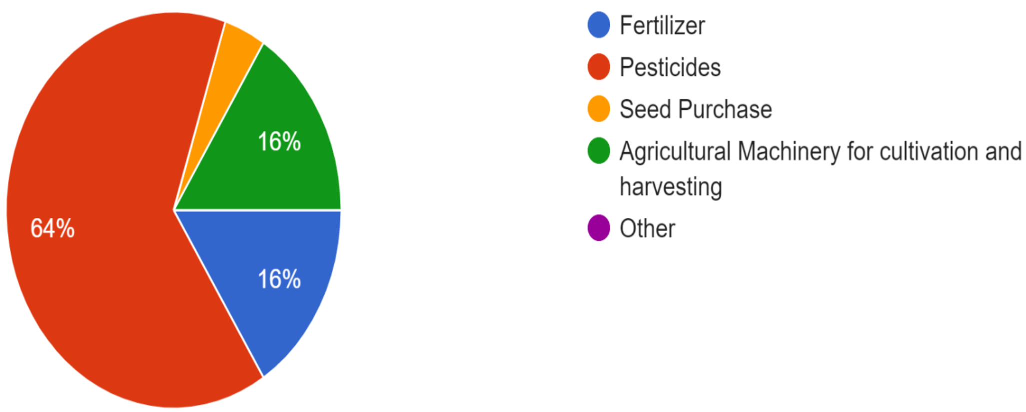




Years of Crop Production
25 responses
32%
8%
24%
32%
10 years and below
11-20
21-30
31-40
41-50
51 years and above

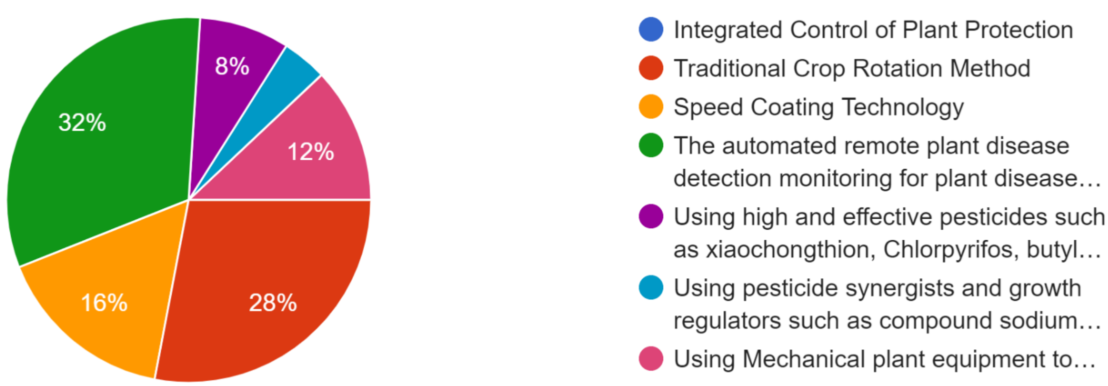
Your farms largest agricultural expenses

25 responses



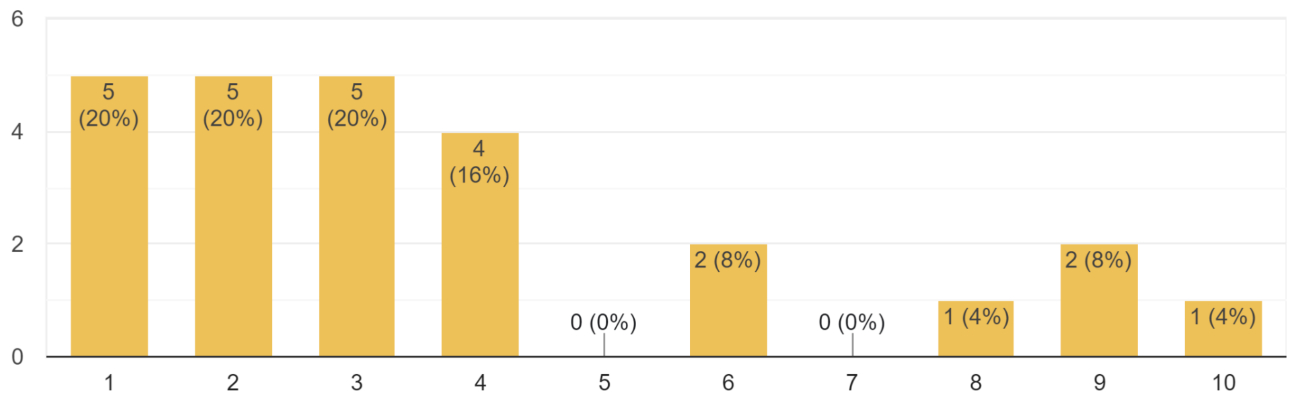
Please tick the options of farming technologies you have heard of for plant disease and insect pests management

25 responses



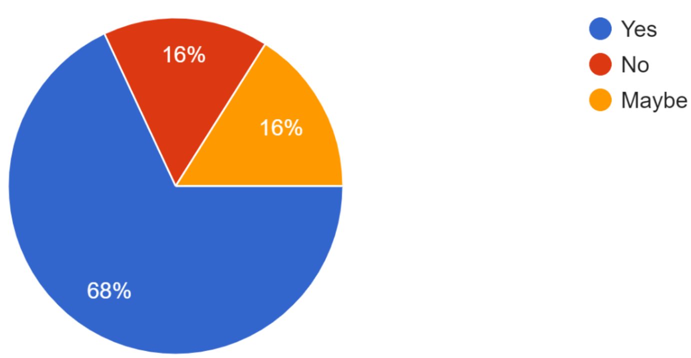
On a scale of 1-10 how well do you know about the automated plant disease diagnosis using a trained machine learning model?

25 responses



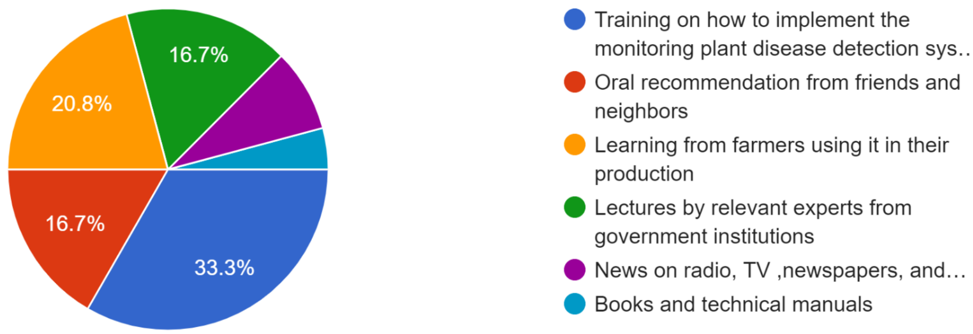
Do you need such technologies

25 responses



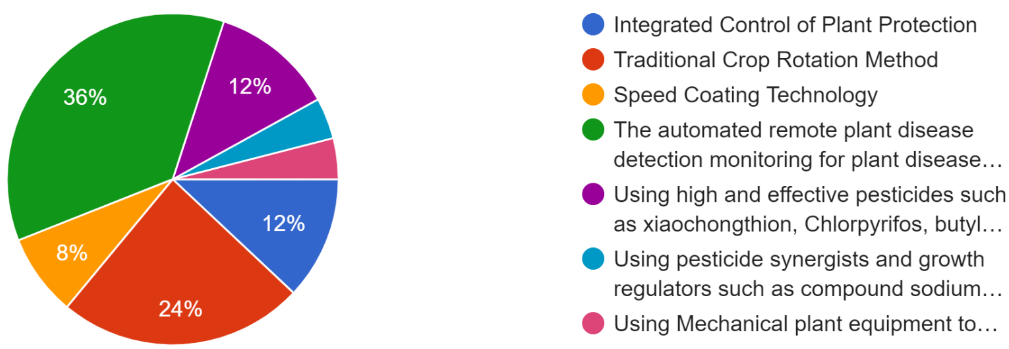
Through which medium do you think it will be best to learn about the technology?

24 responses



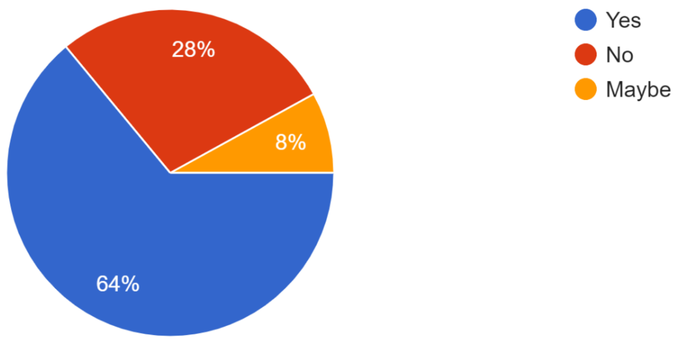
What technology would you like to implement in your farm?

25 responses



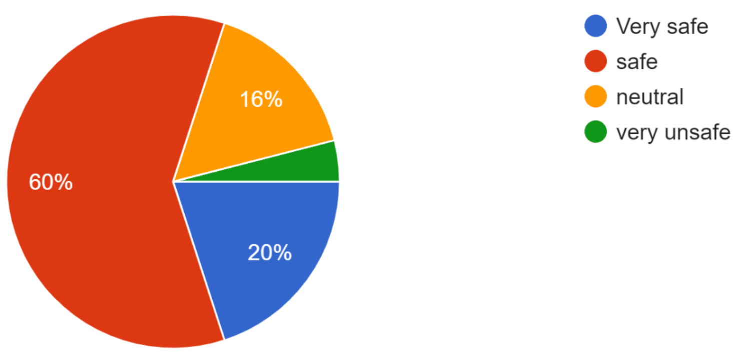
Are you satisfied with the programs implemented by the Ministry of Food and Agriculture in helping increase crop yield production to promote Food Security?

25 responses



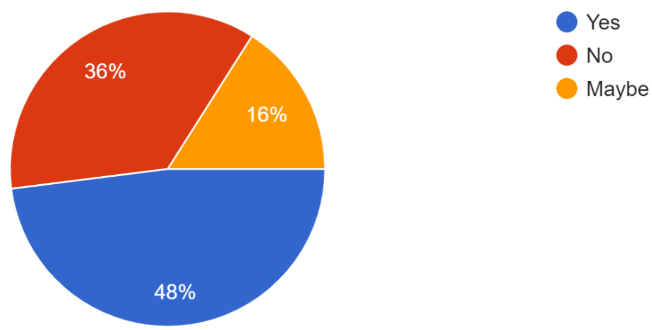
Do you think the impact of adopting automated plant disease detection on crops is safe?

25 responses



Do you think the risk of adopting automated plant disease detection on crops is large?

25 responses



What are some government policies you know about such as government subsidies for pesticides, training programs for plant disease and pest control ...rtilizer technology. Are you satisfied? If not, why?

25 responses