SMART ASSISTANCE FOR THE FLORICULTURE INDUSTRY

Gamage M.G.U.D.

(IT19169736)

B.Sc. (Hons) Degree in Information Technology
Specialized in Data Science

Department of Information Technology

Sri Lanka Institute of Information Technology
Sri Lanka

September 2023

SMART ASSISTANCE FOR THE FLORICULTURE INDUSTRY

2023-133

Final Project Thesis

Gamage M.G.U.D.

(IT19169736)

Dr. Sanvitha Kasthuriarachchi
Dr. Lakmini Abeywardhana

The dissertation was submitted in partial fulfilment of the requirements for the B.Sc. (Honors) degree in Information Technology specialized in Data Science

Department of Information Technology

Sri Lanka Institute of Information Technology
Sri Lanka

September 2023

DECLARATION

I declare that this is my own work, and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	**Student ID**	**Signature**
Gamage M.G.U.D.	IT19169736	

The above candidate has carried out research for the bachelor's degree Dissertation under my supervision.

Signature of the Supervisor: Date:

……………………………. ..……………………….

Dr. Sanvitha Kasthuriarachchi

Abstract

Sri Lanka has emerged as a pioneering force in floriculture production, with growers establishing both small and large-scale businesses across the island. However, these growers face significant challenges stemming from a lack of smart solutions and inefficient technology. Presently, they rely on manual processes for critical tasks such as monitoring plant growth, maintaining export standards, predicting demand and supply, selecting crops, and identifying varieties. Moreover, newcomers entering the industry encounter difficulties related to plant selection, environmental factors, resource management, and plant identification. This underlines the crucial need for a comprehensive smart assistance system to educate and support growers in the field. While numerous research studies have gone into the development of crop recommendation systems and plant identification tools utilizing advanced technologies like the Internet of Things, machine learning, data analytics, deep neural networks, and artificial intelligence, there remains a significant gap in the availability of a plant recommendation system tailored specifically for cut flower crops. Furthermore, there is an absence of identification systems for the Philodendron plant. The proposed solution materializes as a cross-platform mobile application, with a particular emphasis on assisting newcomers to the floriculture industry. This application features a cut flower crop selection system that harnesses an open weather API to gather and process climate data based on the grower's location. Moreover, it takes into consideration infrastructure and resource-related factors when offering recommendations. In addition to these functionalities, the application integrates a variety identification system for the popular Sri Lankan plant, the Philodendron, which represents a novel addition to the industry's toolkit. The successful development and deployment of this system will require expertise in various subjects and technologies, including machine learning, deep learning, transfer learning, image classification, Flask, and React Native. The proposed system holds the potential to become a powerful tool for the Sri Lankan floriculture industry, empowering growers to make informed decisions and enhance their business operations.

Keywords: floriculture, cut flower, plant selection, beginners, weather data, resource, plant identification, philodendron, machine learning, deep learning, image classification, transfer learning, Sri Lanka, mobile application

ACKNOWLEDGEMENT

I want to take this opportunity to thank everyone who was incredibly important in assisting and supporting me as I completed the project component successfully.

First, I would like to express my sincere gratitude to the CDAP team and the research panel for their diligence and commitment during the course. Their dedication to giving us a worthwhile educational experience and guaranteeing the effective completion of our research project has been invaluable. Our project supervisor, Dr. Sanvitha Kasthuriarachchi, and co-supervisor, Dr. Lakmini Abeywardhana, deserve special gratitude for their steady leadership and helpful suggestions, which guided our research in the proper direction.

I would also like to sincerely thank our external supervisor, Mr. Anandathissa Illeperuma, for his invaluable advice during the project. I would like to thank the entire Omega Green Pvt. Ltd. staff for their assistance in data collection, knowledge, and priceless advice. I appreciate the constant assistance and cooperation of my research group colleagues throughout this project.

Lastly, I am deeply thankful for the unwavering support and encouragement from my family and friends, whose belief in me has been a constant source of motivation for successfully completing this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
API	Application Programming Interface
CNN	Convolutional Neural Network
DL	Deep Learning
HTTP	Hypertext Transfer Protocol
IDE	Integrated development environment
IoT	Internet of Things

JSON	JavaScript Object Notation
KNN	K-Nearest Neighbors
ML	Machine Learning
NMT	Neural Machine Translation
RGB	Red, Green, Blue
RL	Reinforcement Learning
RNN	Recurrent Neural Network

LIST OF APPENDICES

1 INTRODUCTION

The floriculture industry is regarded as a high-income-producing agricultural enterprise that has the potential to be used as a tool for Sri Lanka's social and economic development. [1] Growers are establishing small and large-scale enterprises all across the island, but many face problems since there is a severe shortage of smart solutions and outdated technology. Monitoring plant growth, maintaining export standards, forecasting demand and supply, selecting plants to grow, and identifying varieties of floriculture plants are important factors to consider in the floriculture sector. However, most of these tasks are now carried out manually, using knowledge from the past.

The floriculture sector presents a variety of challenges for beginners. They are unsure of what plants to grow considering their current situation. Some important natural factors that affect plant growth are temperature, water availability, and humidity. Depending on where they're located at the moment, these variables will alter. The performance of the floriculture producer is influenced by a number of elements, including financial stability, personnel pool, and space availability. There is no effective system in place for educating beginners about these issues. Additionally, people might misidentify plants, particularly those used in floriculture like Philodendron, which come in a variety and can be difficult to distinguish at first glance. There isn't a simple or effective approach to learning this knowledge. Therefore, Sri Lanka's floriculture business is in severe need of a smart assistance system.

To select appropriate crops, numerous agricultural research projects have been carried out. IoT is used in studies [2] and [3] to gather climate data and predict the best crop using ML algorithms. In order to determine a crop factor and advise farmers to cultivate the crop with the best score, research [4] uses a data analytic algorithm. RNNs are used in research [5] to forecast seasonal weather. Systems for identifying varieties have also been created in earlier research. In research [6] , a deep neural network from Google's Inception V3 is used to create a system for flower recognition. An Anthurium grower's system for variety identification has been presented in research [7] . Deep learning vision transformers have been used in research [8] to develop a Grapevine identification system. Another study [9] was done to determine the Apricot varieties based on the traits of the leaves and a KNN classifier.

According to the literature review, there is currently no system for recommending plants for cut flowers, and the crop recommendation systems that have been created so far depend on weather data collected by IoT to determine which crops are most suited for a certain region. Based on the grower's location, the proposed system uses an OpenWeather API to gather and analyze climate data. It also takes infrastructure- and resource-related issues into account when making predictions. The proposed system uses a dataset that contains information about popular cut flowers grown in Sri Lanka, including anthuriums, orchids, and many others. Additionally, there is no philodendron or variety identification system. The proposed system is based on a dataset of well-known philodendron varieties from the floriculture industry. To successfully complete the proposed system, a variety of topics and technologies must be learned. Examples include transfer learning, ML, DL, image classification, etc.

1.1 Background & Literature Survey

Sri Lanka offers a wide range of outstanding floriculture products and is a global leader in this field. Sri Lanka made USD 16 million from exports of floriculture in 2021. [10] Growers are starting both small- and large-scale businesses all across the island, but many are having difficulties due to the major lack of intelligent solutions and the outdated technology currently in use. Key factors to consider in the floriculture sector involve monitoring plant growth, maintaining export standards, forecasting demand and supply, selecting plants, and identifying varieties. However, depending on previous knowledge, the majority of those tasks are now done manually.

Beginners in the floriculture industry face problems such as trouble selecting plants to grow, difficulty identifying plant varieties, a lack of education, and a lack of resources. It is important to choose a plant that is appropriate for the environment and infrastructure because there are numerous elements that might affect how well a plant grows. A study by Kristin L. Getter in 2015 that examined the effects of average daily temperature on four floriculture crops—Geranium, Petunia, Marigold, and Pineapple mint—showed that temperature is a significant factor impacting these plants. According to the study's findings, the growth of flowers and plants is significantly influenced by the average daily temperature [11] . According to a study by Leiv M. Mortensen [12], who studied four floriculture crops, including Begonia, Chrysanthemum, Poinsettia, and Blossfeldiana, air humidity has an impact on the maintenance of flower crops as well. The highest plant quality and fewest plants were often produced at the lowest humidity levels. Despite the fact that the results differed by species and variety, the key discovery was that air humidity had a significant influence on sustaining life during growth. Research [13] conducted in 2022 to investigate important elements associated with plant development indicated that light, water, and nutrition are additional major components that affect plant growth. Understanding these elements is crucial when managing plants to increase production and meet demand. Several plant types' health and behavior are influenced by the availability of water. Land and space availability are crucial when considering infrastructure factors impacting the growth of plants. Advanced plant nurseries frequently require a lot of space because crops can be raised for several years in containers or in the ground before being sold. Nurseries that grow adult trees in straight rows for transplant or sale require a minimum of one hectare of land in order to be profitable [14]. The availability of skilled labor is one of the industry's key advantages, according to the industry capabilities report from 2021. Universities offer courses in floriculture to develop skilled workers for industry [10]. According to a 2017 study by Padmini, S. M. P. C., and Kodagoda, T. D. [1], this sector faces significant financial difficulties when compared to others like agriculture. In order to avoid future problems, newcomers to the business should take seriously the above environmental and infrastructure concerns. Otherwise, banks and other financial institutions may be unwilling to grant financial assistance to this industry.

In addition, beginners in the industry lack expertise and have a limited understanding of plant identification. Decorative foliage for export is produced in Sri Lanka primarily through out-grower systems and cooperatives of farmers. Additionally, Philodendrons are frequently utilized as tissue-cultured plants and decorative foliage [10]. There are over 450 different species of Philodendron [15]. According to a study [16], Philodendron plants can exhibit a broad variety of differences in their leaves, making it challenging to distinguish between varieties. In the agricultural industry, identification of plant types and varietals is crucial because of the vast diversity of plants found within various crop species, according to a study [17]. There are no procedures in place to let individuals know this. As a result, newcomers encounter various challenges and obstacles, which demotivate them. Since Philodendron has such broad appeal and would benefit thousands of floriculture growers, it is vital to establish an effective and trustworthy system for variety identification.

Several agricultural studies were carried out in order to determine an appropriate crop. IoT is used in research [2] and [3] to collect climatic data like humidity, temperature, and rainfall. And ML algorithms are used to forecast the best crop. Crop scoring is a data analytics algorithm used in research [4]. Based on rainfall, soil type, cropping month, and location, a crop factor has to be calculated. The crop with the greatest score should be grown, as the grower is instructed. According to the study [5], RNNs are combined with random forest classification methods, estimated weather characteristics, and soil parameters to predict seasonal weather. The system for identifying flowers has been developed through research [6]. Rose, Daisy, Tulip, Dandelion, and Sunflower flower varietals are classified using Google's Inception V3 deep neural network. A mobile application for growers of Anthurium has been demonstrated in research [7]. The identification of Anthurium variants makes use of NMT, RL, CNN, and ML technologies. In a study [8], DL vision transformer architecture was utilized to detect different Grapevine kinds using leaf centered RGB pictures. For Apricot leaves, feature extraction was carried out in research [9]. The KNN classifier was used to classify the acquired feature vector.

1.2 Research Problem

For many individuals who wish to discover the possibility of growing and selling plants and flowers, the floriculture sector has emerged as an attractive economic opportunity. The floriculture industry in Sri Lanka currently employs ineffective technology. Less new technology is being used, and less effort is invested in research and development [10]. Plant growth monitoring, maintaining export standards, predicting demand and supply, selecting plants, and identifying varieties are important areas to investigate in the floriculture sector. The majority of these duties are currently, however, carried out manually with the use of past knowledge. The business specifically presents several difficulties, especially for newcomers with limited knowledge and expertise.

Determining the correct crops to grow is one of the major obstacles facing newcomers to the floriculture industry. This is a key decision that could significantly affect the success or failure of a business. The growth and development of a plant can be influenced by a variety of environmental conditions, including temperature, humidity, and water. The choice of plants can also be influenced by the floriculture grower's location. For example, plants that thrive in a hot, humid climate could struggle to grow in a region with lower temperatures. Consequently, choosing the appropriate plant for a specific site is a difficult process, especially for beginners. Beginners with no prior knowledge, however, may be unsure how to take these factors into consideration when choosing the ideal plant to cultivate. Figure 1-1 illustrates how beginner growers choose their plants based on a survey that was conducted.

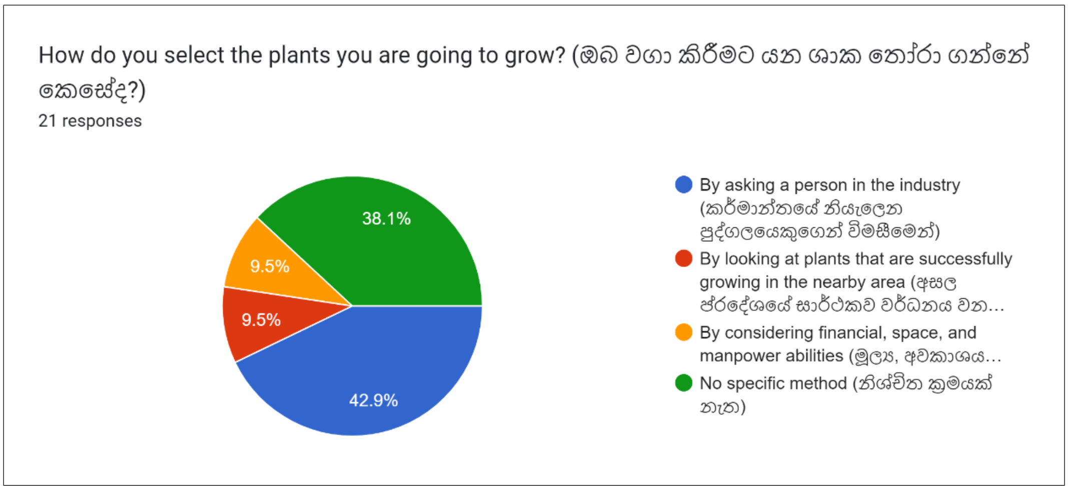


Figure 1-1 Survey to find how beginners select plants

The financial stability, pool of workers, and space availability of the floriculture grower are other elements that influence its performance. The availability of water sources has an impact on a plant nursery's ability to succeed. Therefore, while choosing a plant, resource-related aspects should also be taken into account. While growing some plants may require a lot of space, time, and money, growing others may be simple and require fewer resources. Growers can run into issues when expanding their nurseries if the plants grow well. They could not have the resources or available space to expand the business [1]. But there isn't enough machinery in place to choose which plants to grow based on those factors. If the plant does not thrive, the grower will be in difficulty, lose money and effort, and waste resources. The suggested system uses the OpenWeather API and ML to determine the best plant to grow in an area by taking into consideration various environmental aspects like temperature and humidity. Additionally, the proposed system takes into account the availability of water supply, space, and money. The survey found that an automated solution would be useful for this. (Figure 1-2)

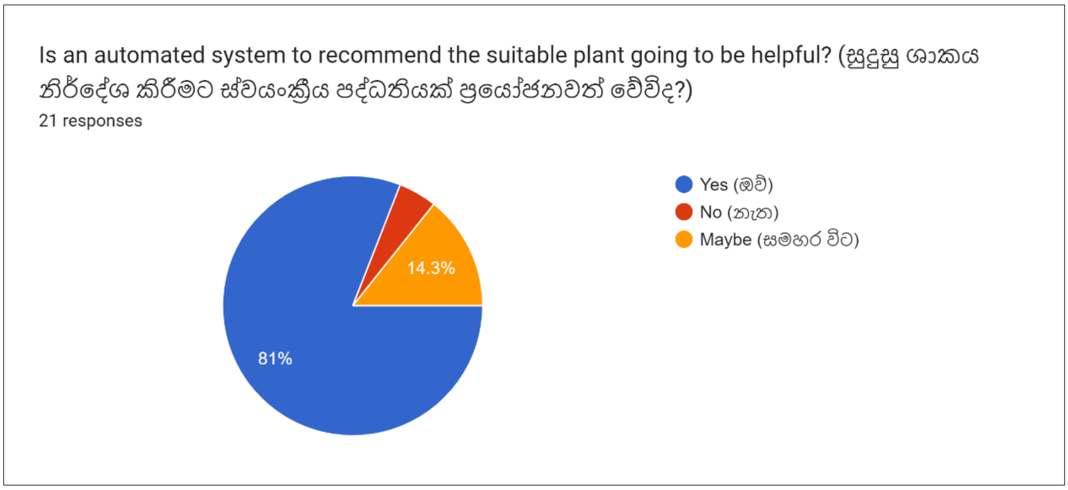


Figure 1-2 Survey to identify if an automated system is useful

The identification of different varieties of plants is another problem in the floriculture industry. In Sri Lanka, there are many diverse plant species, making it challenging to distinguish between them. This problem mostly affects plants with cut foliage because there are numerous varieties of one plant type. Dieffenbachia, Cordyline, Philodendron, Scindapsus, and Monstera make up the majority of Sri Lanka's exported cut foliage [18]. There are numerous varieties of Philodendron, such as Congo green, Birkin, Majesty, Black cardinal, Moonlight, Sun red, and many more. Because of this, it could be challenging for a first-timer to tell the difference between varieties. Furthermore, there are no proper and easily accessible sources to identify Sri Lanka-based plants. Beginners, therefore, encounter difficulty in this situation. Figure 1-3 shows that most of them lack a proper and dependable approach.

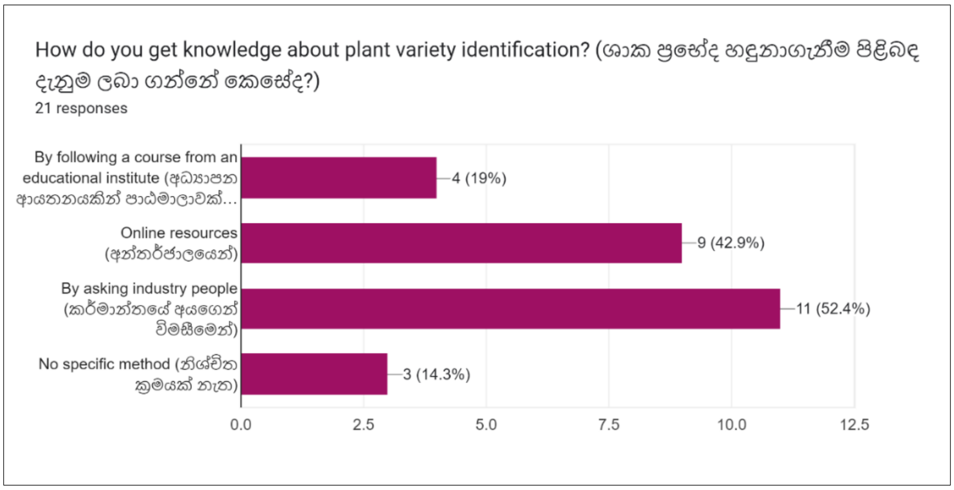


Figure 1-3 Survey on the knowledge sources of beginners

One of the numerous harms of inaccurately identifying plant varieties might be improper maintenance, which impairs a plant's growth and health. As a result of customer dissatisfaction, the market might experience financial losses. Additionally, it can make it more difficult to control diseases and pests, resulting in the overuse of pesticides and greater costs. Inaccurate identification can also frustrate newcomers and keep them from learning about and appreciating the diverse plants, like Philodendrons. The survival of these kinds depends on correct identification in order to prevent these possible issues. For those new to floriculture, an automated Philodendron variety identification system utilizing deep learning is crucial as a first step in creating a variety identification system. It not only simplifies the process of cultivating and enjoying the various Philodendron varieties but also ensures accurate sales, promotes understanding of these plants, and helps with the care of different varieties. The survey results are shown in Figure 1-4, and they support the value of an identifying system.

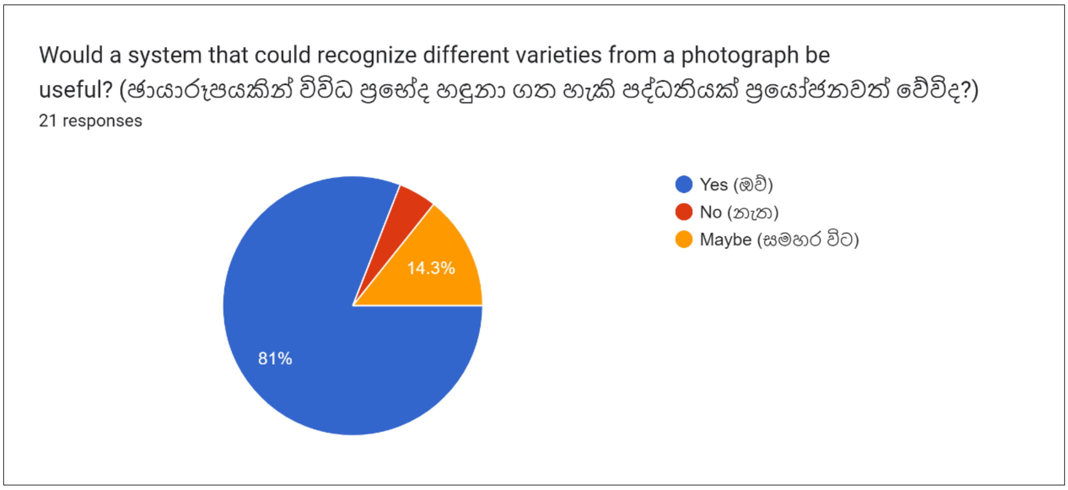


Figure 1-4 Survey on the usefulness of a variety identification system

By doing so, the proposed system would be beneficial to beginners in the floriculture industry, making it easier.

1.3 Research Gap

The following is a description of related research done locally as well as internationally.

Research [2] is about an IoT and ML-based agriculture system that will assist farmers and agriculturists predict crop yields. The goal of this study was to predict an efficient crop that may be cultivated in a specific field area and result in a high crop yield. Predictions are made using metrological data such as temperature, humidity, soil moisture, and so on. Metrological data is gathered via an IoT-based sensor system. The crop is predicted using the ML algorithm.

In [3], the user, would be able to use this system to get correct advice on which plant will thrive depending on various factors such as humidity, pH, and rainfall. IoT is used to measure these factors. To train the dataset and make the forecast, ML algorithms were utilized.

In [4], Crop scoring is a data analytic algorithm that has been used in research. Rainfall and soil type were acquired based on region, and farmers provided cropping month and location. The crop factor is then calculated using a mathematical model that assigns values to variables and adds them. The farmer is advised to grow the crop with the highest score.

[5] RNN is being used for seasonal weather predicting. Crop selection models incorporate soil factors such as soil type, pH, fertility, and water holding capacity. The soil and estimated weather characteristics are combined to select appropriate crops for land. The random forest classification technique is used to classify appropriate crops.

[6] Describes a DL-based strategy for flower identification and classification. The study's image samples include images of a variety of flowers, including Tulips, Roses, Daisies, Sunflowers and Dandelions. In pursuit of greater accuracy in less time, Google's Inception V3 deep convolutional neural network's final layers have been re - trained for classifying on the ornamental plants data set.

[7] Introduces a mobile application for detecting 5 different types of Anthurium plants, managing plant care activities, diagnosing 3 diseases, offering safety measures, forecasting 5 pests, helping planters, and providing a way for planters to locate a market, assess export quality, and predict the most popular kind. NMT, RL, CNN, image processing, and ML technologies are used.

In [8], researchers have utilized advanced deep learning vision transformers for the purpose of distinguishing between 12 distinct Grapevine varieties. The dataset is made of RGB images of leaves of the grape vines.

[9] Explains how leaf features can be used to identify different Apricot varieties. A dataset of leaf photos from 10 different Apricot types was created as part of the study. The background was removed from the leaves during the segmentation process. Apricot leaf characteristics were used to determine structural aspects. then the k-nearest neighbor classifier was used to classify the acquired feature vector.

Based on the studies conducted, it has been observed that there are various crop recommendation systems available for agriculture on an international level. However, when it comes to floriculture, there is no existing plant recommendation system to assist growers in selecting the best cut flower plants for their specific needs. The crop recommendation systems that have been developed so far utilize weather factors to predict the most suitable crops for a particular area. However, these systems usually rely on IoT devices to measure weather factors accurately.

The proposed system, on the other hand, makes use of OpenWeather API to collect and process climate data based on the grower's location. Furthermore, the previously developed crop recommendation systems have not focused on the grower's infrastructure and resource-related factors, such as land size, cost, and water availability, when predicting the ideal plant. However, the proposed system considers these factors when providing recommendations.

Although there have been studies conducted on flower and plant variety identification using DL techniques, no solution has been developed to identify Philodendron varieties. The proposed system can identify six Philodendron varieties, which are mainly used in cut foliage fields in floriculture. They are Congo green, Birkin, Majesty, Black cardinal, Moonlight, and Sun red. By simply entering an image of the plant, the system can identify the variety. And the mobile app includes smart validation using CNN. It can identify whether the image is within the six varieties mentioned or not.

Table 1-1 compares the proposed system with past studies.

Table 1-1 Past research

Reference / **Features**	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	Proposed system
Obtain the climate data without hardware devices	-	-	-	-	-	-	-	-	✓
Predict the best cut flower plant to grow considering weather and infrastructure	-	-	-	-	-	-	-	-	✓
Identify varieties of Philodendron	-	-	-	-	-	-	-	-	✓
Assess if the plant falls within any of the six Philodendron varieties.	-	-	-	-	-	-	-	-	✓
Mobile application	-	✓	-	-	-	✓	-	-	✓

1.4 Research Objectives

1.4.1 Main objective

Cut flower recommendation and Philodendron variety identification for floriculture industry beginners.

1.4.2 Specific objectives

- Obtain the climate data without hardware devices like sensors based on the grower's location.
- Predict the best cut flower plant to grow out of eight cut flower crops, considering weather and infrastructure factors.
- Differentiate the main six varieties of Philodendron using an image of the plant.

- Validate the user's input to determine if the plant falls within any of the six Philodendron varieties.
- Cross-platform mobile application development

1.5 Project Requirements

1.5.1 Functional requirements

- System should be able to obtain the climate data forecast based on the location.
- System should be able to get infrastructure data from the user.
- System should be able to predict the best cut flower plant to grow.
- System should be able to capture a photo of a plant using the mobile phone camera.
- System should be able to validate the user input image.
- System should be able to identify the Philodendron variety.

1.5.2 Non-functional requirements

- Reliable: The system should be dependable and trustworthy, and it should consistently produce accurate and consistent results.
- Fast: The system should be designed to perform tasks quickly and efficiently, with minimal delays or wait times.
- Accurate: The system should be designed to provide accurate and precise results, with a high level of data integrity
- User-friendly: The system should be easy to use and navigate, with a simple and intuitive user interface with clear instructions

1.5.3 System requirements

- Google Colab is used for machine learning and deep learning development.
- TensorFlow, Keras, Scikit-Learn, and OpenCV libraries are used for model development.
- PyCharm as an IDE and Flask as a framework are used for backend API development.

- Postman is used for API testing.
- Visual Studio Code as an IDE and React Native as a framework are used for front-end development.
- EXPO is used for app development.

1.5.4 User requirements

- The user should be able to enter their location to the app to obtain weather data for their area.
- The user should be able to enter details about the available infrastructure they have.
- The user should be able to capture a picture of a Philodendron plant they want to identify with their smartphone camera and have the app identify the plant's variety name.

2 METHODOLOGY

For the purpose of providing smart support to the floriculture sector, the ‘Plant Pal’ smartphone application has been created. Its features encompass growth monitoring, maintenance of export standards for affected plants, supply and demand forecasting, and assistance for beginners. Figure 2-1 illustrates the overall system architecture.

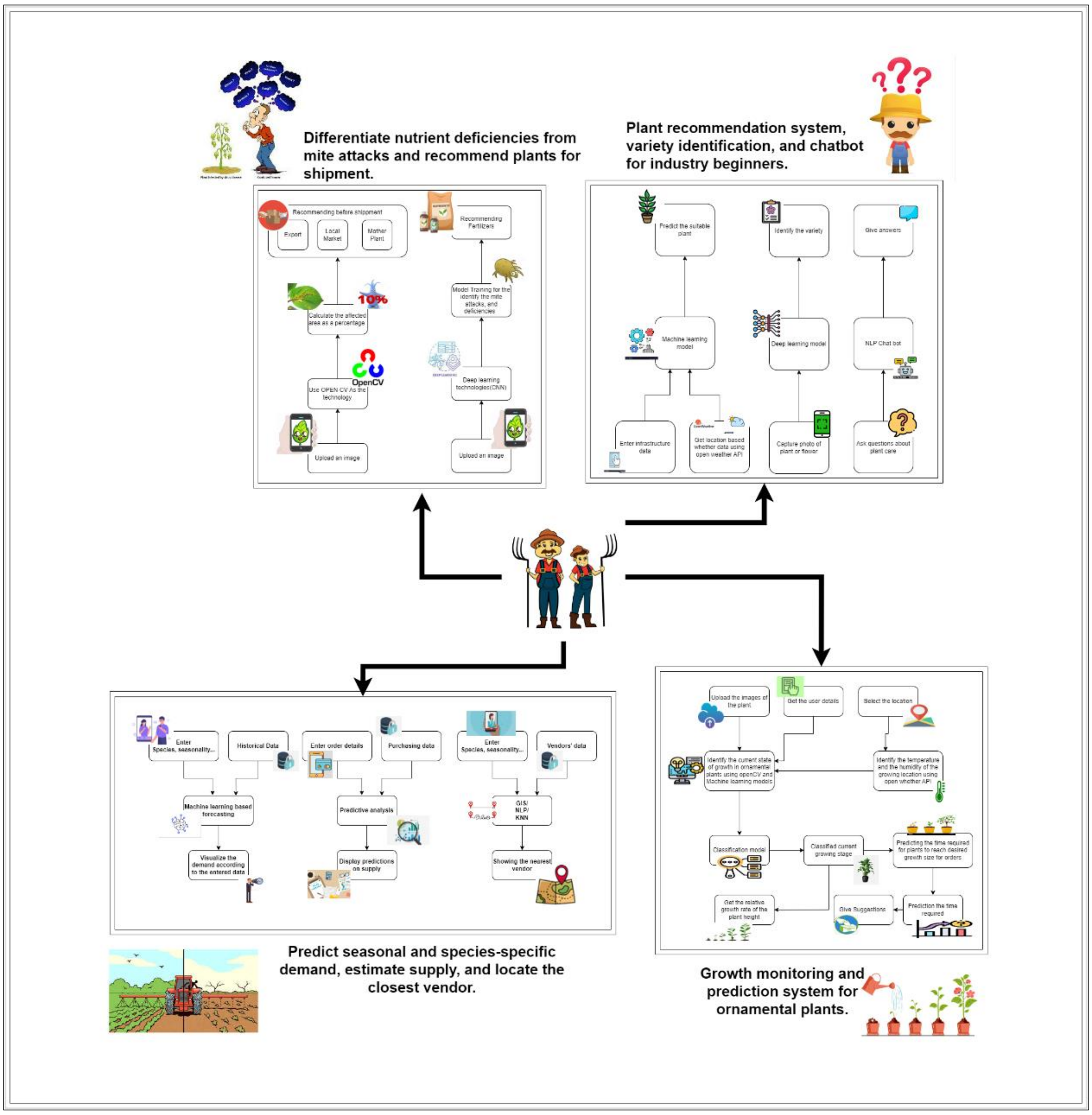


Figure 2-1 Overall system architecture

The individual component of the Plant Pal mobile app named ‘For Beginners’ is implemented as two subparts: ‘Cut flower Advisor’, the recommendation system, and ‘Philo Variety Finder’, an identification system. Figure 2-2 explains the individual component of the system architecture. Including plant recommendation system and variety identification system.

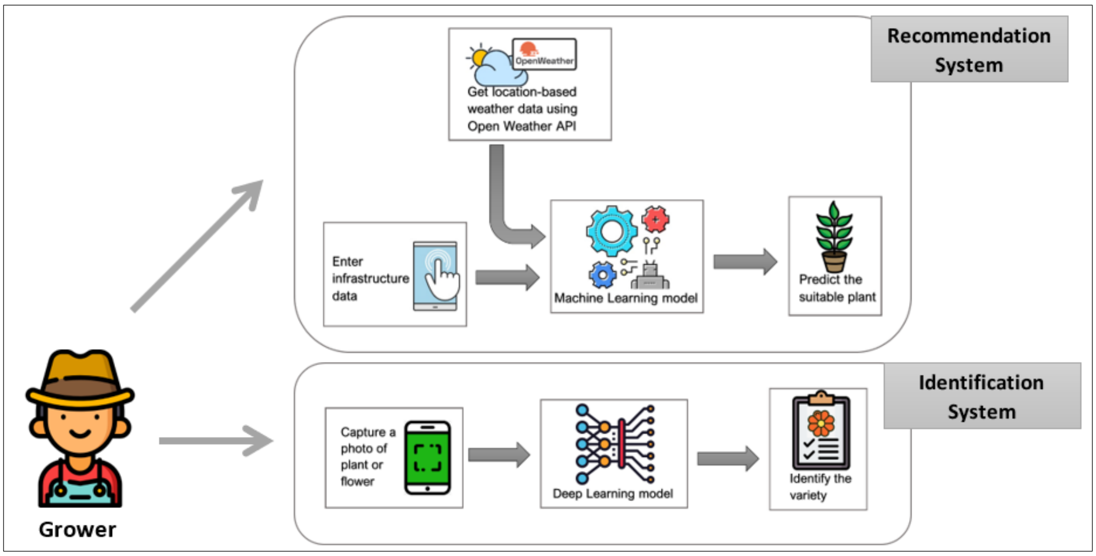


Figure 2-2 Individual system diagram

Table 2-1 summarizes the technologies, algorithms, architectures, and techniques which are used in the implementation of the proposed system.

Table 2-1 Summary of tools and technologies used

Technologies	**Algorithms & architectures**	**Techniques**
React Native Python Tensorflow Keras scikit-learn Flask OpenCV Open Weather API	Supervised learning CNN	Data preprocessing Feature engineering Transfer learning Data augmentation Hyperparameter tuning

The suggested system would be developed in accordance with the Agile software development process shown in Figure 2-3.

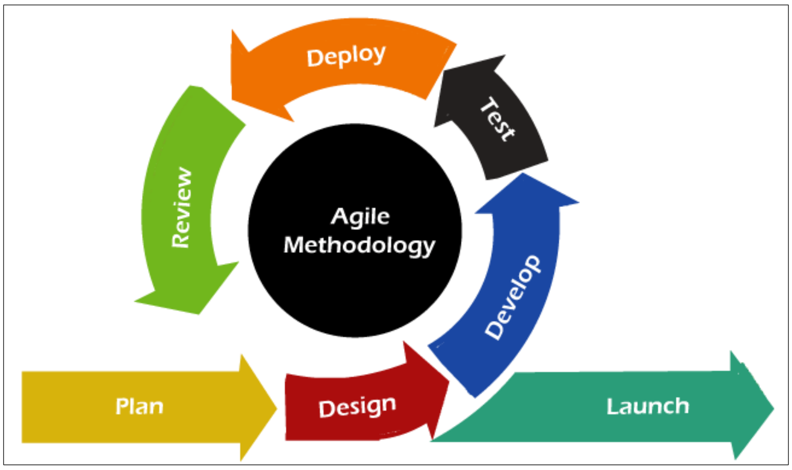


Figure 2-3 Agile methodology

2.1 Commercialization & Business Plan

Plant Pal, a smartphone application developed for the floriculture sector, provides two appealing features, Cut flower Advisor and Philo Variety Finder, accessible to both beginners and ambitious business owners. It functions well on iOS and Android platforms and also serves as a flexible backend API.

Plant Pal's business strategy includes subscription offers, strategic partnerships, data monetization, and ongoing development, with a focus on assisting interns and beginners in floriculture and encouraging initial entrepreneurs. Plant Pal seeks to establish an audience in the market, becoming a necessary tool for sector enthusiasts while assuring profitable business development. This is accomplished by offering insightful data and encouraging user loyalty.

2.2 Backend Implementation

2.2.1 Crop recommendation system

- **Data collection**

The weather and resource-related data about the best-grown cut flower nurseries among island wide cut flower growers were collected from growers including Omega Green Pvt. Ltd. The data set included temperature and relative humidity in the area, water liters consumed in a year, land or net house size in square feet, number of plants in the nursery, initial net house and land preparation costs, initial planting costs, annual labor costs, annual maintenance costs for fertilizers and pesticides, and cut flower yields in a year with the relevant cut flower crop name. Eight cut flower crops are covered: Roses, Carnations, Alstroemeria, Anthurium, Gerbera, Orchids, Lilies, and Chrysanthemums. There were over 2300 records in the dataset.

- **Data preprocessing**

The gathered data was separated into features and labels. The features were the weather and resource-related data, and the label was the cut flower crop name. The rows with null or missing values were removed. Then the dataset was split into training and testing, with a test size of 33%.

- **Exploratory data analysis**

To display the number of samples in each class, a plot was taken. In Figure 2-4, it can be seen that the data is balanced among 8 classes.

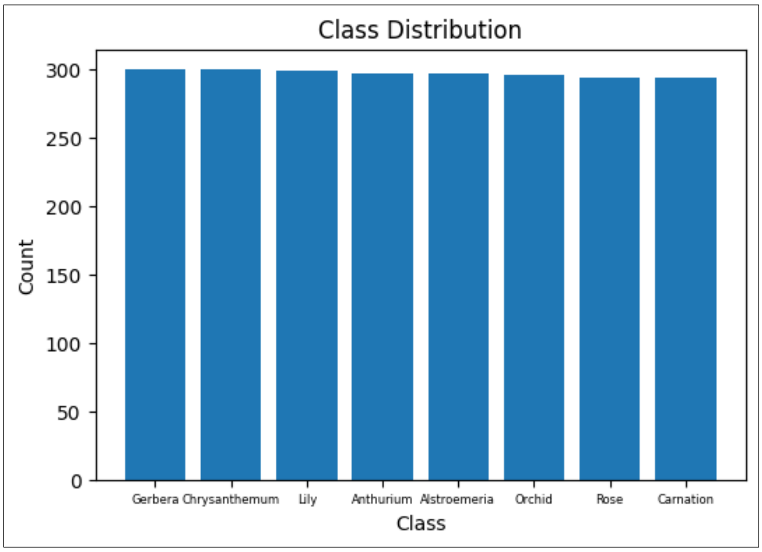


Figure 2-4 Cut flower class distribution

To visualize the distribution of data for each feature, histograms were obtained as in Figure 2-5.

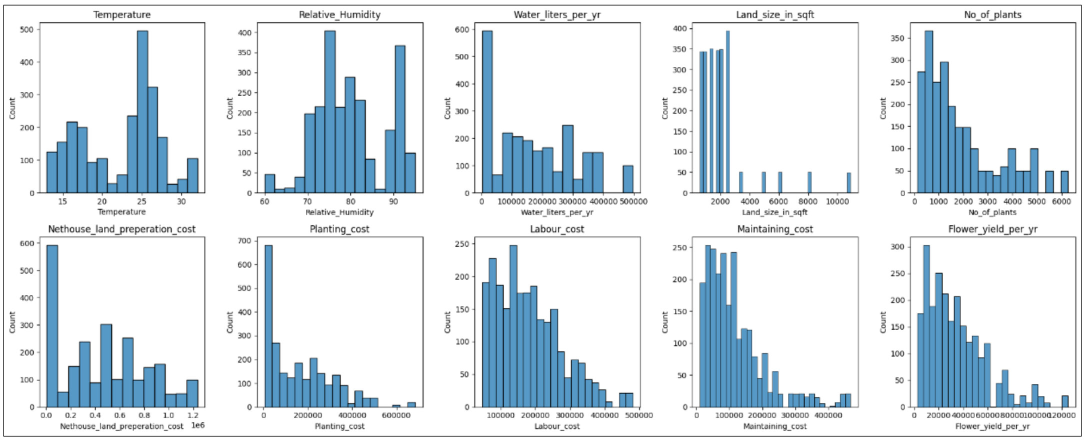


Figure 2-5 Histograms for each feature

To discover more about the relationships between the dataset's various attributes, a correlation matrix was created, with the colors representing the correlations' strength and direction. In Figure 2-6, the obtained confusion matrix is displayed.

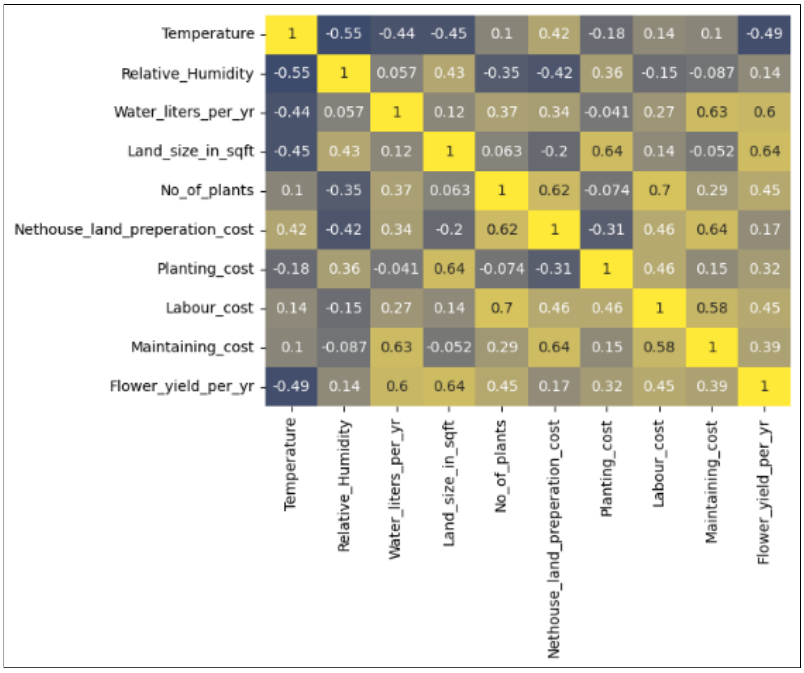


Figure 2-6 Cut flower confusion matrix

Temperature, as an important factor in crop growth, was studied in relation to the cut flower crop. Figure 2-7 illustrates the graph.

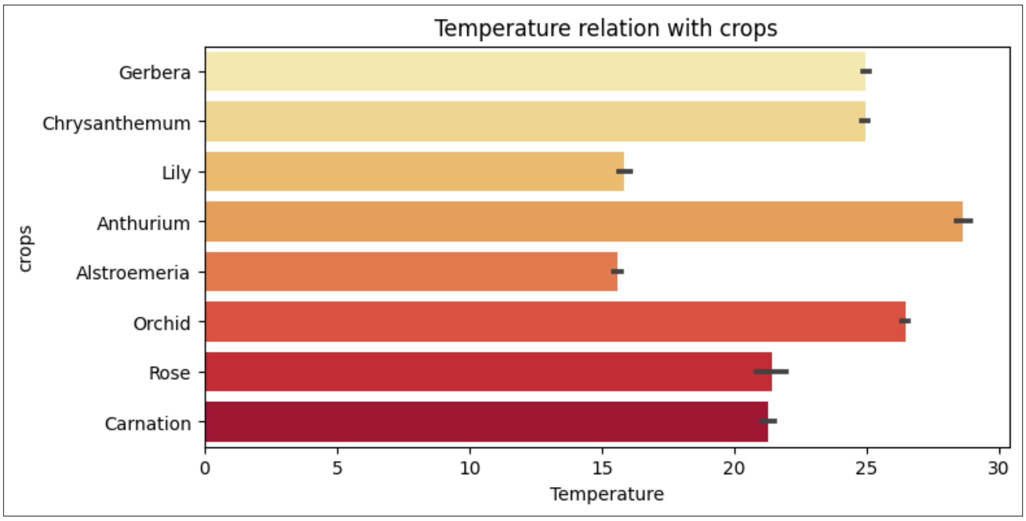


Figure 2-7 Temperature relation with crop

- **Feature selection**

'SelectKBest' method from scikit-learn was used to perform feature selection based on F-scores, which measure the difference in means between classes relative to the variance within each class. F-score is a statistical measure that is often used in analysis of variance (ANOVA). This plot provides a visual representation of the importance of each feature for classification as shown in Figure 2-8. Features with higher F-scores contribute more to distinguishing between classes. Then the top 9 features were selected for classification. Which are:

- Temperature
- Relative Humidity
- Water liters per year
- Land size in square feet
- Number of plants
- Nethouse or land preparation cost
- Planting cost
- Maintaining cost
- Flower yield per year

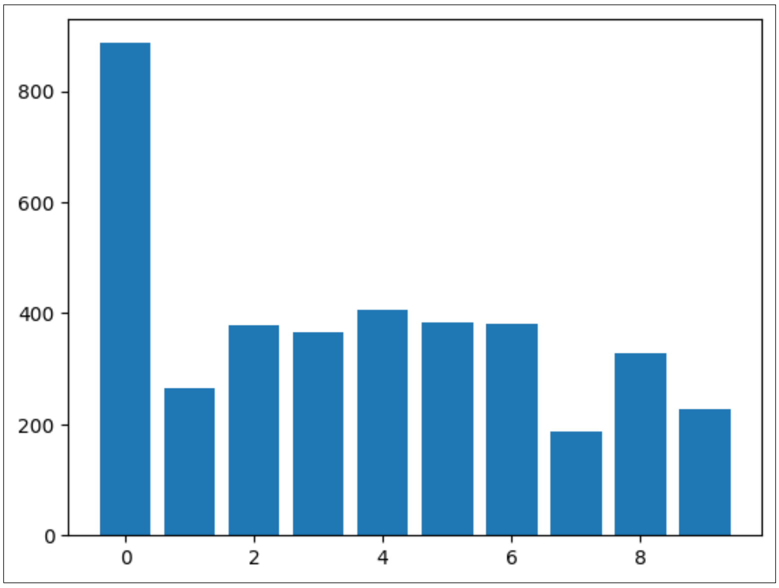


Figure 2-8 Feature importance

- **Model comparison**

To suggest the best cut flower crop to grow out of 8 classes, 3 ML algorithms were considered best known for multiclass classification: Decision Tree, Naive Bayes, and Gradient Boosting. The Decision Tree classifier was selected because it is a good starting point for understanding the data's structure. Naive Bayes performs well even with relatively small datasets. Gradient Boosting's Ensemble method combines the predictions of multiple weak learners. To evaluate these 3 models, Stratified K-Fold cross-validation with 3 folds was used. Then, the mean accuracy and standard deviation of the accuracy scores for each model were obtained. The Gradient Boosting model gave an exceptionally high mean accuracy of approximately 99.94%. The extremely low standard deviation (0.000888) showed that the model's performance is highly consistent across folds, indicating a strong and stable performance. So, the Gradient Boosting model was selected as the best model. Table 2-2 displays the model accuracies and standard deviation values.

Table 2-2 ML algorithm comparison

ML Algorithm	Mean accuracy	Standard deviation
Decision Tree	0.994978	0.005821
Naive Bayes	0.902018	0.011549
Gradient Boosting	0.999372	0.000888

- **Model training**

The selected Gradient Boosting model was trained using the previously selected best 9 features. The number of boosting stages = 50, the learning rate = 0.01, and the maximum depth of each decision tree = 1 were defined as hyperparameters when training. The learning curve was obtained to visualize how the model's accuracy changes as it goes through the boosting stages, and the smoothing applied to the curves will make the trends more noticeable. The x-axis represents the number of boosting stages, and the y-axis represents the accuracy. The learning curves for both the training and test datasets are displayed on the same plot (Figure 2-9).

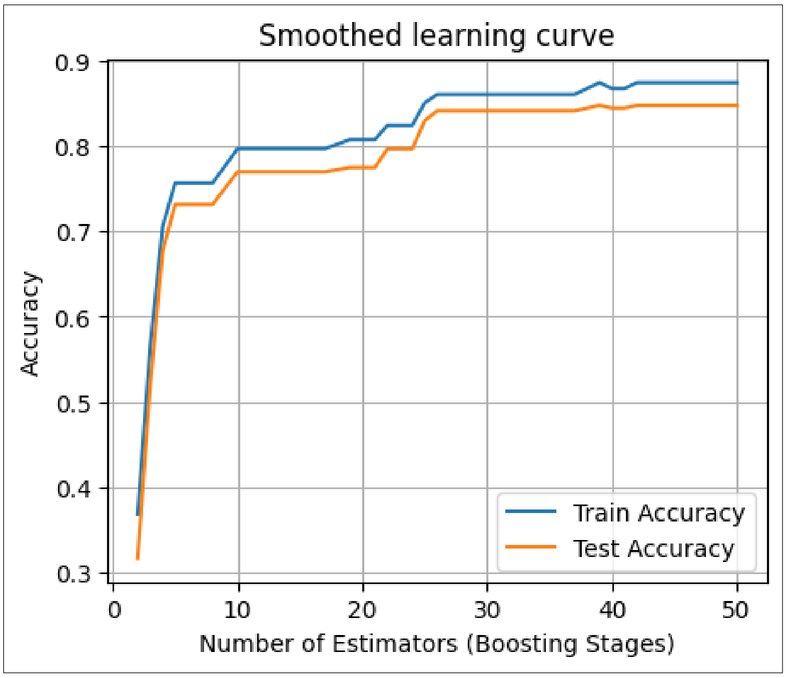


Figure 2-9 Cut flower learning curve before hyperparameter tuning

- **Hyperparameter tuning**

In the hyperparameter tuning process for the Gradient Boosting Classifier, a range of hyperparameters to optimize the model's performance was explored. These included the number of boosting stages (n_estimators), which determines the complexity of the ensemble, the maximum depth of individual trees (max_depth), which controls the depth of each decision tree in the ensemble, and the learning rate (learning_rate), which regulates the step size during optimization. By systematically searching through various combinations of these hyperparameters using scikit-learn's ‘GridSearchCV’, the best set of hyperparameters was identified to enhance the model's accuracy. The best hyperparameters were – learning_rate: 0.1, max_depth: 1, n_estimators: 100. Then, the model was retrained with these hyperparameters. Below are the learning curves after hyperparameter tuning for training and testing. The improvement in model accuracy is visualized in this plot (Figure 2-10).

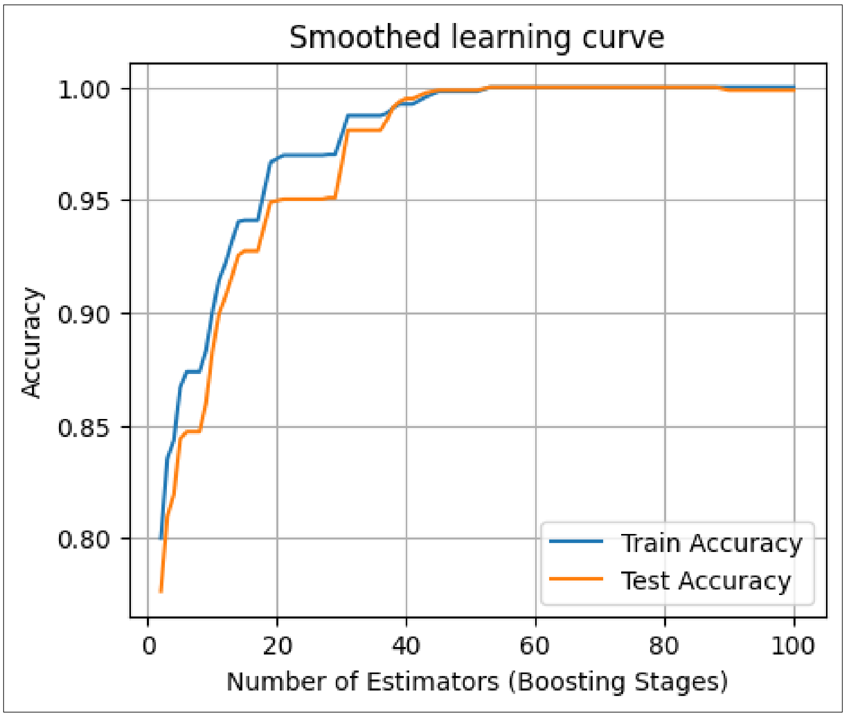


Figure 2-10 Cut flower learning curve after hyperparameter tuning

- **Location based weather data retrieving**

OpenWeather API was used to fetch current weather data, including temperature (in Celsius) and relative humidity, for a given city. It creates an API call, sends a request, handles the response, and shows the weather data as in Figure 2-11.

```
http://api.openweathermap.org/data/2.5/weather?q=Kandy&appid=71d1945ed6fb6c764f9334a8b0c1dc4a
Temperature: 21.31 °C
Humidity: 97%
```

Figure 2-11 OpenWeather API output

- **API development**

The API is designed to receive HTTP POST requests, process the data contained in those requests, and return responses in JSON format. Flask is used as the web framework, Python is the programming language, and PyCharm is the IDE used to develop the API. First, a pre-trained Gradient Boosting machine learning model is loaded using the Pickle module.

When predicting user inputs, the city name is extracted from the JSON request data. It is sent to the OpenWeather API, which retrieves the temperature in Celsius and the relative humidity. Additional user inputs related to water usage, land size, number of plants, several costs, and flower yield are retrieved from the JSON request data. All collected data (weather parameters, user inputs) is combined into a single input array. This array is then converted into a NumPy array for prediction. The loaded ML model predicts the cut flower crop. The prediction result is packaged into a JSON response and returned to the user.

2.2.2 Variety identification system

- **Data collection**

Mainly, two datasets were created: the Philodendron classification dataset and the Binary classification dataset.

Philodendron classification dataset: for the classification of 6 Philodendron varieties, 603 images were collected from Omega Green (Pvt) Ltd. The dataset included images of Congo green, Black cardinal, Moonlight, Sun red, Majesty, and Birkin.

Binary classification dataset: for the purpose of validating if the image is a Philodendron within this scope or not, a binary classification dataset was created. For that, another 606 images were collected; these images included random objects extracted from the ImageNet dataset and also images of Philodendron plants, which are out of the 6 varieties mentioned. This class is named ‘non-philo’. The previously collected 603 Philodendron images of 6 varieties were also included as one class named ‘philo’ for this binary classification dataset. The total dataset size is 1209.

- **Data preprocessing**

The first step of preprocessing was resizing images to a fixed size of 224x224 pixels using OpenCV to ensure all images had the same dimensions. The next step was scaling the pixel values of the images by dividing them by 255. This operation was carried out to scale the pixel values from the range [0, 255] to [0, 1], as it helps with convergence during training. The data were split into training, testing, and validation. For training the model, 75% of the data was allotted; for validation during training, 12.5% of the data was allotted; and to evaluate the model's performance after training, 12.5% of the data was allotted.

- **Exploratory data analysis**

Image counts for each class in the Philodendron classification dataset were visualized to check the data balance (Figure 2-12).

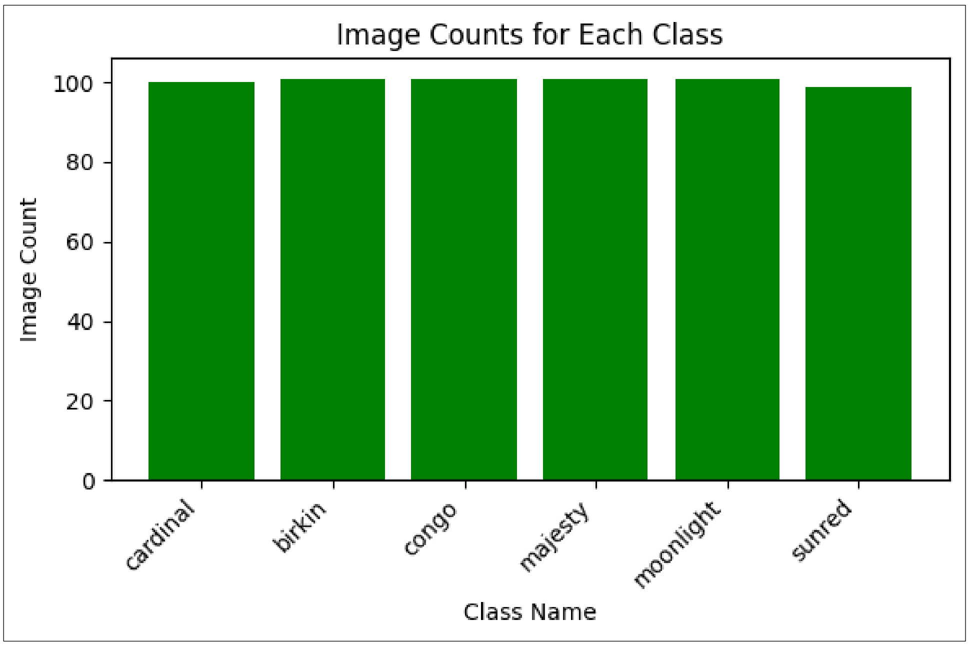


Figure 2-12 Philodendron dataset class distribution

Image counts for each class in the Binary classification dataset were visualized to check the data balance (Figure 2-13).

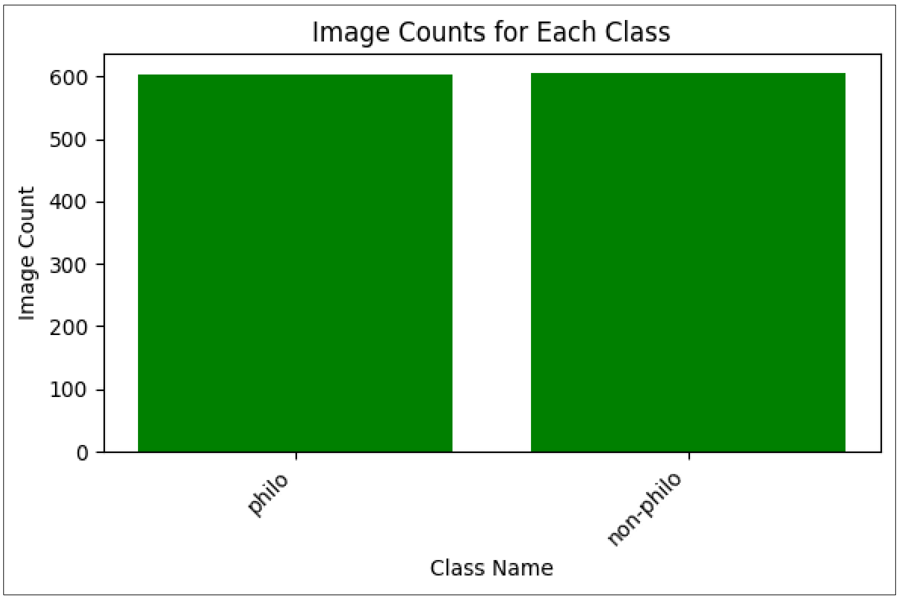


Figure 2-13 Binary dataset class distribution

The distribution of train, test, and validation samples across classes was visualized in the Philodendron classification dataset (Figure 2-14).

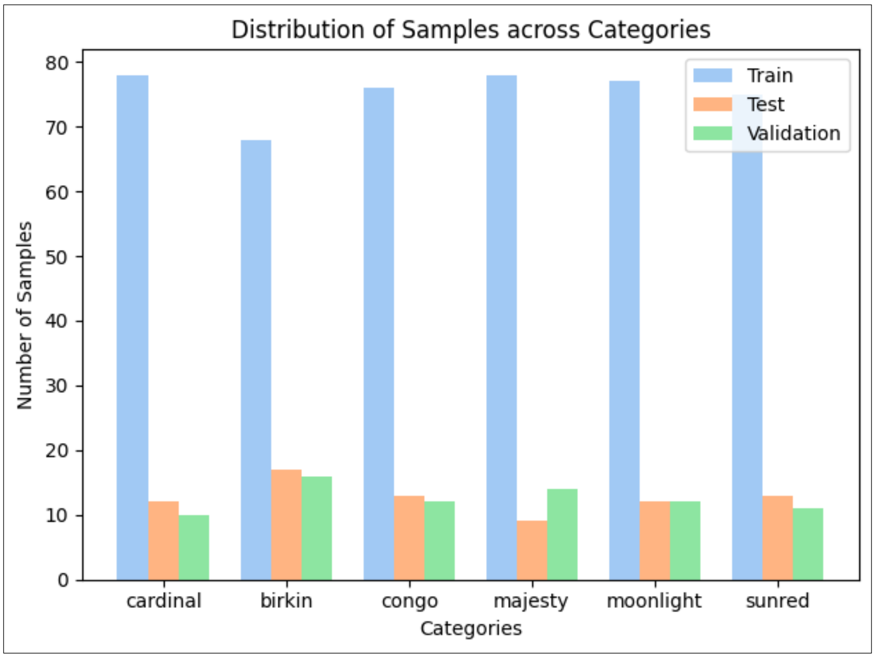


Figure 2-14 Philodendron dataset sample distribution

The distribution of train, test, and validation samples across classes was visualized in the binary classification dataset (Figure 2-15).

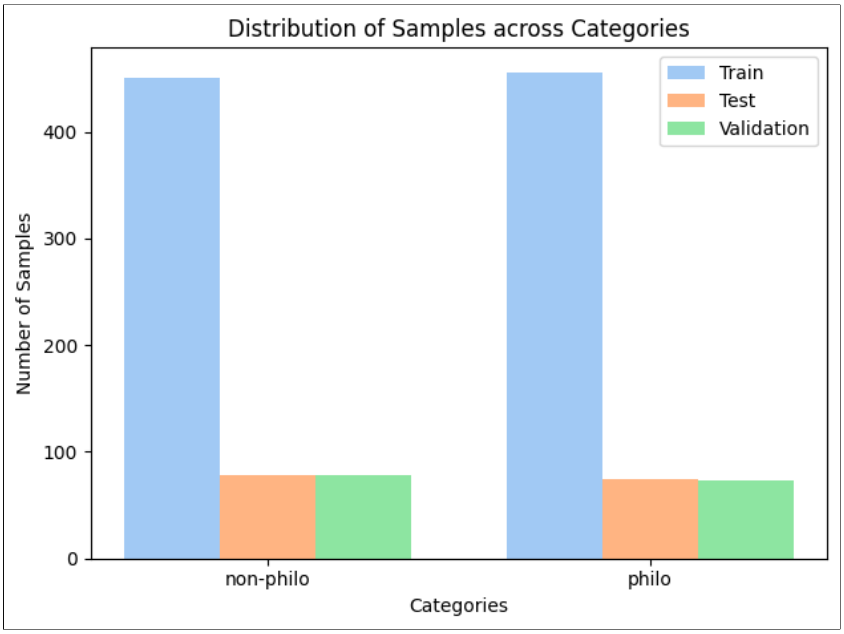


Figure 2-15 Binary dataset sample distribution

- **Data augmentation**

Using the Keras API of TensorFlow, data augmentation techniques were used to 2 already-prepared image datasets. By making images to random transformations including flips, rotations, and zooms, data augmentation improves the training dataset. This helps improve a model's capacity to generalize and function well on unseen data.

Using the RandomFlip (horizontal_and_vertical) layer, images were randomly flipped both horizontally and vertically, changing their orientation. By randomly rotating images a maximum of 0.2 radians in any direction, the RandomRotation (0.2) layer increased the variety of image orientations. The RandomZoom (0.2) layer was used to randomly enlarge or reduce images by up to 20%, which helps the model adapt to changes in object scale. Figure 2-16 illustrates how a sample image appears after data augmentation.

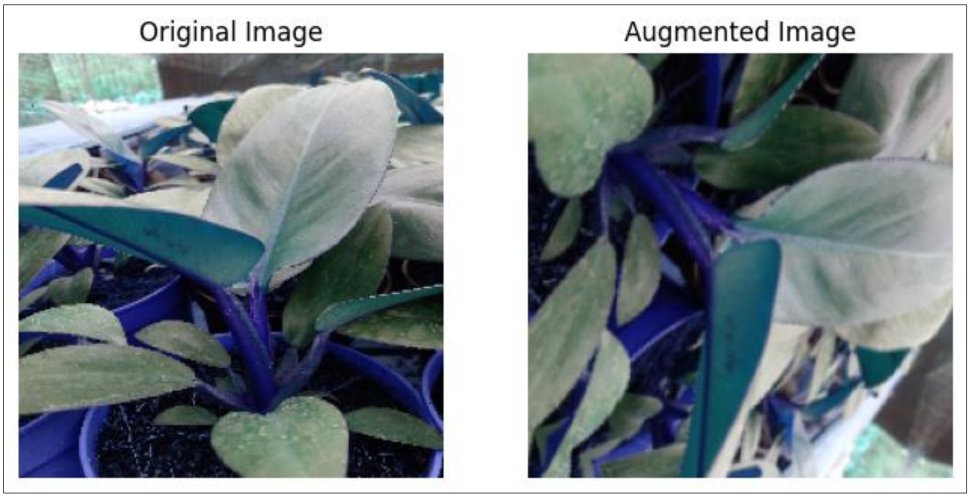


Figure 2-16 Data augmentation

- **CNN Model architectures for Philodendron classification**

a. Transfer learning model architecture

Transfer learning is a technique in deep learning where a model trained on one task is leveraged to perform better on a related but different task. It involves using the knowledge learned from one dataset to improve the performance of a model on another dataset. Transfer learning is particularly advantageous for small datasets because they typically suffer from overfitting. So, in this Philodendron variety classification into 6 classes, transfer learning was applied.

The CNN model architecture used was a sequential neural network built using TensorFlow's Keras API. The data augmentation layer was the first layer of the model. As the next layer of the CNN model, a feature extractor model was used. A pre-trained MobileNetV2 model from TensorFlow Hub was used as the feature extractor. The weights were frozen in this pre-trained model to prevent them from being updated during training. MobileNetV2 is pre-trained on a large dataset and can extract useful features from input images. MobileNet is a family of CNN architectures designed for efficient on-device vision applications. They are known for their small model sizes and low computational power consumption.

MobileNetV2 is an improved version of the MobileNet architecture with better performance.

After the feature extractor, a dense layer with 256 units was added. This layer was used to learn high-level features from the representations generated by the pre-trained model. The activation function used was ReLU (Rectified Linear Unit), which introduces non-linearity into the model. Regularization is applied using L2 regularization with a strength of 0.01. By penalizing huge weights in the network, it aids in preventing overfitting. Following the first dense layer, a dropout layer was included. Dropout randomly deactivates a fraction of neurons during each training iteration, which helps prevent overfitting by adding a form of regularization. The dropout rate in this layer was set to 0.3, which means, on average, 30% of the neurons will be turned off during training. The final layer in the model was another dense layer with the number of units equal to the number of philodendron classes, which is 6. The activation function used in this layer was Softmax, which converts the model's output into class probabilities. Similar to the previous dense layer, L2 regularization with a strength of 0.01 was added.

b. Custom model architecture

A custom CNN model was developed using TensorFlow's Keras API. The model was started with an input layer of shape (224, 224, 3). The model consisted of multiple pairs of convolutional layers, each followed by a max-pooling layer. Two succeeding convolutional layers with 16, 32, 64, 128, and 256 filters, respectively, were added. ReLU activation functions and same padding were used by these layers. After each pair of convolutional layers, a max-pooling layer with a (2, 2) pooling window was used to compress the feature maps. Batch normalization was applied after the second set of convolutional layers and the third set of convolutional layers. The inputs to each layer were normalized, helping with training stability. Dropout layers with dropout rates of 0.5, 0.3, and 0.2 were included after the final convolutional layers and between dense layers. After the last max-pooling layer, a flatten layer was used to convert the 2D feature maps into a 1D vector, which could then be fed into the dense layers. The model was included with two dense layers with 256 and 128 units, respectively, and the ReLU activation function was used. These layers were also included before batch normalization and dropout layers. The output layer was a dense layer with a number of units equal to the number of Philodendron varieties, which is 6. The Softmax activation function was used, and the predicted class probabilities were the output. L2 regularization with a strength of 0.01 was applied to the weights of the dense layers.

- **Model training**

The 2 model architectures were separately trained using the prepared dataset of 603 Philodendron images from 6 classes. The Adam optimizer was used, which is a popular choice for training neural networks. Plots for training and validation accuracy as well as training and validation loss over the epochs of the transfer learning model and the custom model's training are as obtained. Figure 2-17 illustrates transfer learning model and Figure 2-18 illustrates custom model performance.

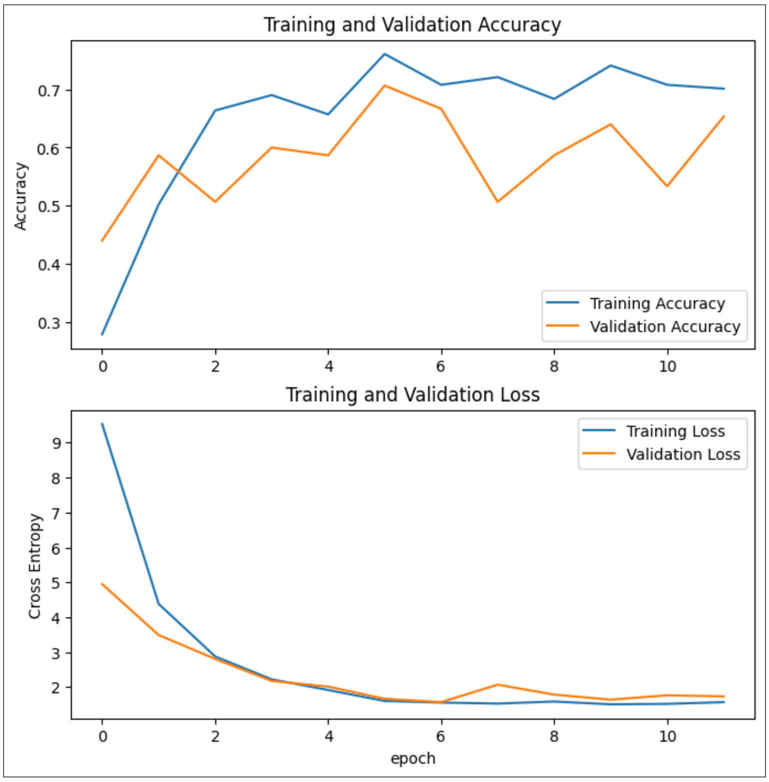


Figure 2-17 Transfer learning model learning curve

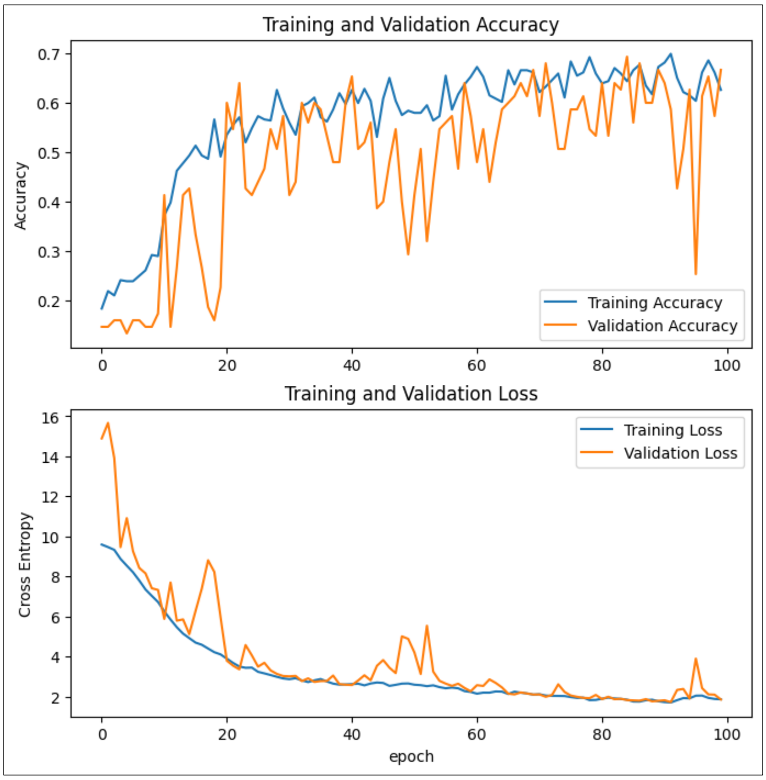


Figure 2-18 Custom model learning curve

- **Evaluating performance**

Evaluation of the performance of the 2 trained models on the test dataset was done using the evaluate method in TensorFlow's Keras API. model made predictions on the test data, calculated the specified evaluation metrics (loss and accuracy), and returned the evaluation results. Figure 2-19 and Figure 2-20 display the transfer learning and custom model performance respectively.

```
model.evaluate(X_test_scaled,y_test)

3/3 [==============================] - 3s 763ms/step - loss: 1.4693 - acc: 0.7632
[1.4693450927734375, 0.7631579041481018]
```

Figure 2-19 Transfer learning model evaluation

```
model.evaluate(X_test_scaled,y_test)

3/3 [==============================] - 4s 1s/step - loss: 1.8148 - acc: 0.6842
[1.8148170709609985, 0.6842105388641357]
```

Figure 2-20 Custom model evaluation

The transfer learning model performed better with an accuracy of 0.76 and a loss of 1.46 than the custom model, so the transfer learning model was selected for further development.

- **Hyperparameter tuning**

Hyperparameter tuning is essential for optimizing model performance, reducing overfitting, and improving efficiency in deep learning. Hyperparameter tuning was set up using the RandomSearch tuner. Random Search randomly samples hyperparameter combinations from the given search space. Different hyperparameter combinations to try were given as 15. Hyperparameters included the number of dense units, dropout rate, regularization strength, learning rate and number of training epochs. After hyperparameter tuning for the transfer learning model, Figure 2-21 shows the best values.

```
Hyperparameters:
num_epochs: 47
dense_units: 128
regularization_1: 1.2423501162154591e-05
dropout_rate: 0.2
regularization_2: 1.001578904444805e-06
learning_rate: 0.001
```

Figure 2-21 Best values after hyperparameter tuning

Using these hyperparameters, the best model was built and trained again. And the learning curve displayed improved accuracy as in Figure 2-22.

Figure 2-22 Learning curve for improved model

- **Early stopping**

During the retraining of the transfer learning model, early stopping was used as a callback to monitor the validation loss. Using this, the training process was terminated if the validation loss did not improve. Additionally, when training is stopped, the callback is configured to return the model's weights to their best state. The best-performing model will be saved as a result.

- **CNN Model architecture for binary classification**

A transfer learning model was developed for classifying user input images as Philodendron within the 6-variety scope or out of scope, which means some random object or a Philodendron other than the 6 varieties. This CNN model was used for validating the user input image before feeding it to the Philodendron classification model.

This binary classification model was also developed using transfer learning with a pre-trained MobileNetV2 feature extractor. The data augmentation layer that was defined earlier was the first layer of the model, and the second layer was the MobileNetV2 feature extractor that was loaded earlier. The next layer was a dense layer with 256 units and a ReLU activation function. L2 regularization with a strength of 0.01 was also included. This layer was used to learn high-level features from the MobileNetV2 feature vector. Then, a dropout layer was utilized to randomly drop 30% of the neurons during training to avoid overfitting. Finally, the output layer of the model had a single unit and a sigmoid activation function. L2 regularization was applied here as well.

The binary classification model was trained using the prepared binary classification dataset. using the optimizer as Adam. Plots for training and validation accuracy as well as training and validation loss over the epochs are illustrated in Figure 2-23.

Figure 2-23 Binary classification learning curve

- **API development**

Python is utilized as the programming language, PyCharm as the IDE, and Flask as the web framework. The developed API is used to identify the variety of Philodendron plants in images submitted via POST requests. It first loads a binary classification model that predicts if the image is a Philodendron within scope or not. If the image is identified as a Philodendron within 6 varieties, it proceeds to identify the specific variety.

The image from the request is preprocessed, including resizing to (224, 224) pixels, converting to an array, pixel value normalization to [0, 1], and batch dimension addition. Then the predicted class number is extracted. If the class number is 0 (not a philodendron within scope or a random object), it returns ‘Unknown’. If the class number is 1 (a Philodendron within scope), it loads a second model, the Philodendron classification model, that predicts the specific variety of the Philodendron. And outputs the predicted variety name.

2.3 Frontend Implementation

The system's frontend is designed as a React Native mobile application with features that allow navigating between several screens. React Native is an open-source framework for creating mobile applications using JavaScript and React. Native mobile apps for iOS and Android can be built with React Native. Visual Studio Code served as the IDE in this front-end implementation. Figure 2-24 displays the home screens of the app. When the ‘For Beginners’ button is tapped, it navigates to the screen where the Cut flower Advisor and Philo Variety Finder are located.

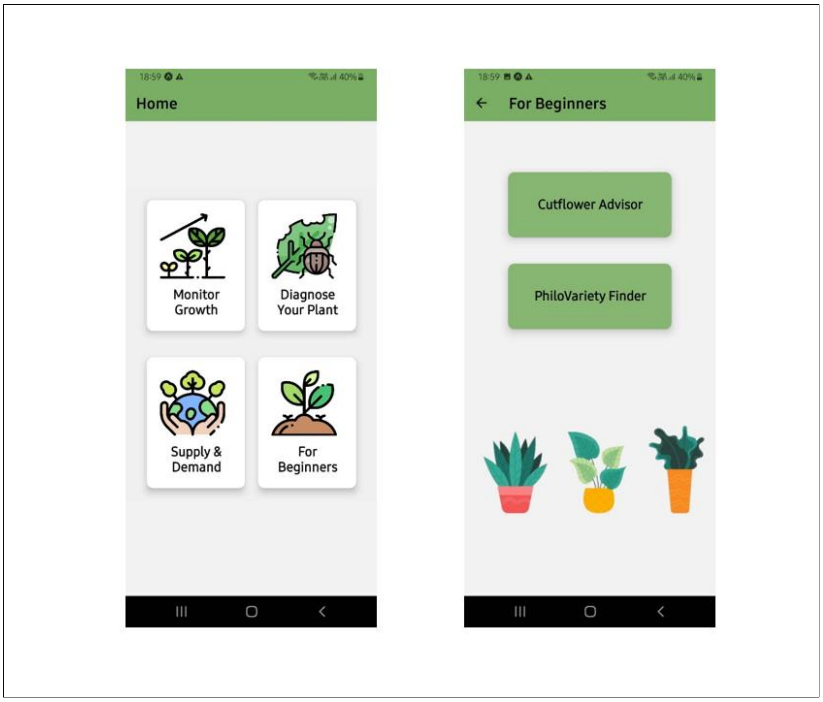


Figure 2-24 Beginners home page

In Cut flower Advisor, the user can input the necessary information and tap ‘Get Suggestion’. It pops up as the best cut flower crop name (Figure 2-25).

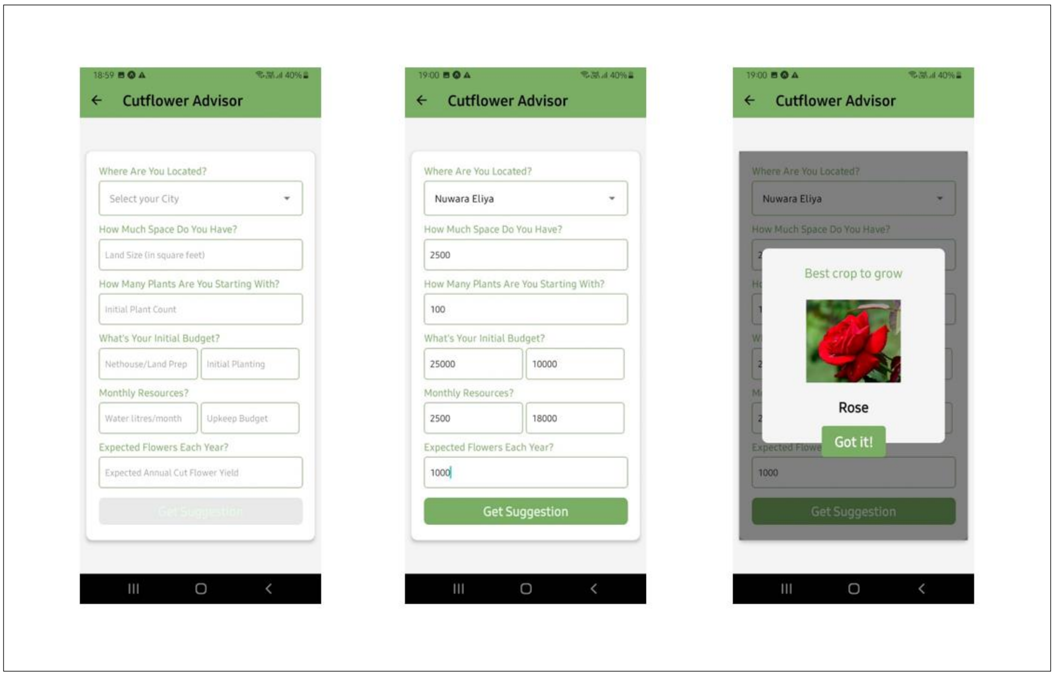


Figure 2-25 Cut flower Advisor screen

In the Philo Variety Finder screen, the user can tap the 'Capture a Photo' button and take a picture of a Philodendron plant. And the 'See the Variety' button shows the predicted Philodendron variety name and a fact about the plant as shown in Figure 2-26.

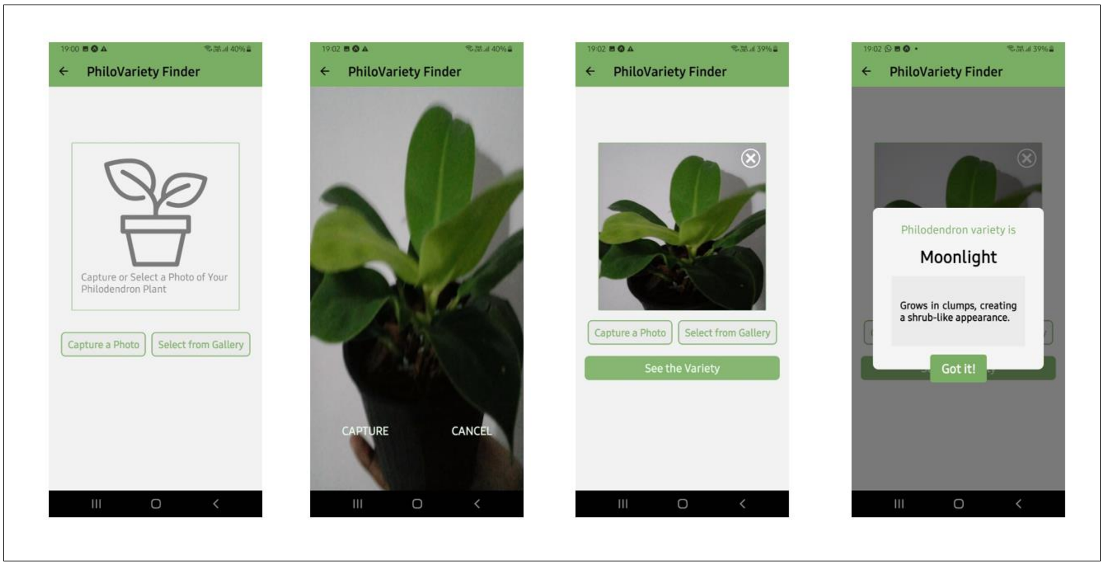


Figure 2-26 Capturing a photo of Philodendron

By tapping 'Select from Gallery', the user can select an image in the gallery and see the variety name (Figure 2-27).

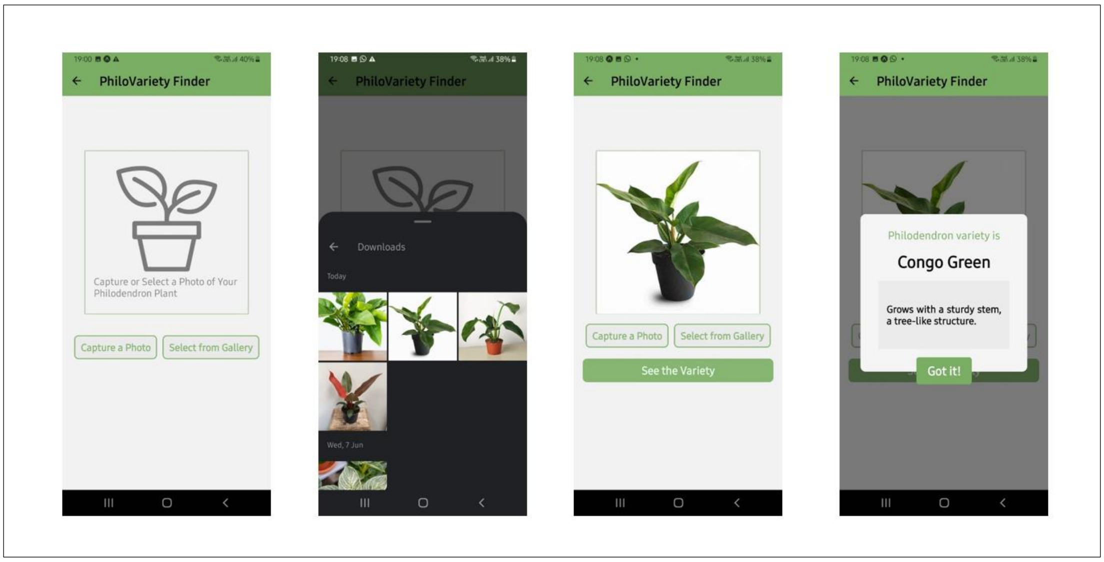


Figure 2-27 Selecting from gallery

If the user accidentally captures or selects an image of a random object or a Philodendron plant out of the 6 varieties, it alerts the user that the image is unknown (Figure 2-28).

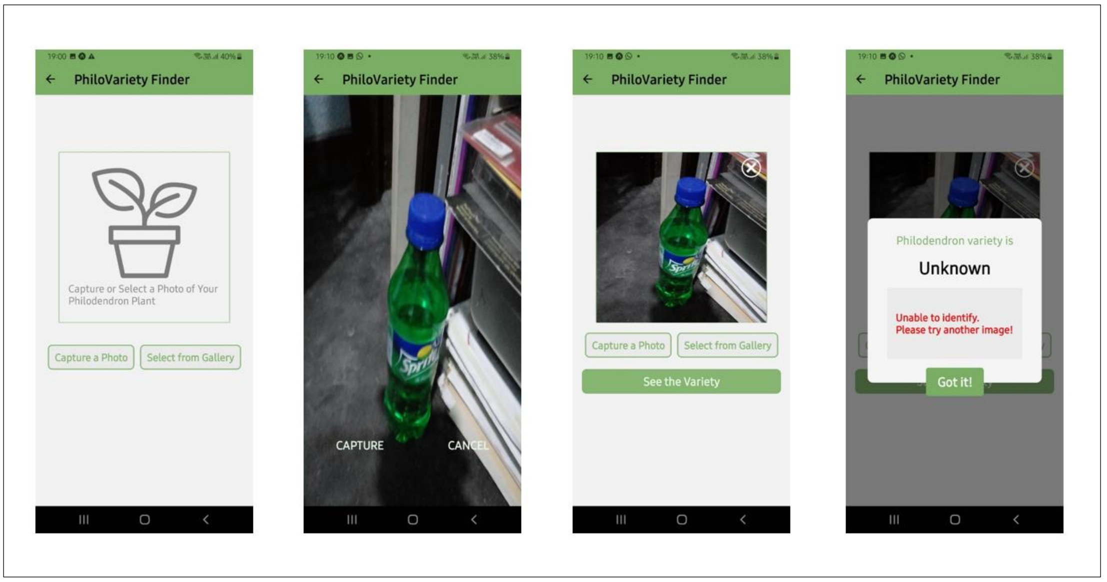


Figure 2-28 Smart validation

3 RESULTS AND DISCUSSION

3.1 Testing and Results

- **Cut flower Advisor**

On the 785 unseen test data records, initial predictions were generated. Afterward, predictions were assessed using the Scikit-Learn module.

a. Accuracy score

This evaluates accuracy by comparing the actual labels with the predicted labels. used to calculate the proportion of samples in the test dataset that were accurately predicted. The Cut flower Advisor model scored a high accuracy score of 99.87%.

b. Confusion matrix

The confusion matrix, which indicates the number of samples correctly or incorrectly identified for each class, delivers useful insights into the model's performance. The count of true positives, true negatives, false positives, and false negatives is displayed. The confusion matrix in Figure 3-1 demonstrates how well the model predicts unseen data.

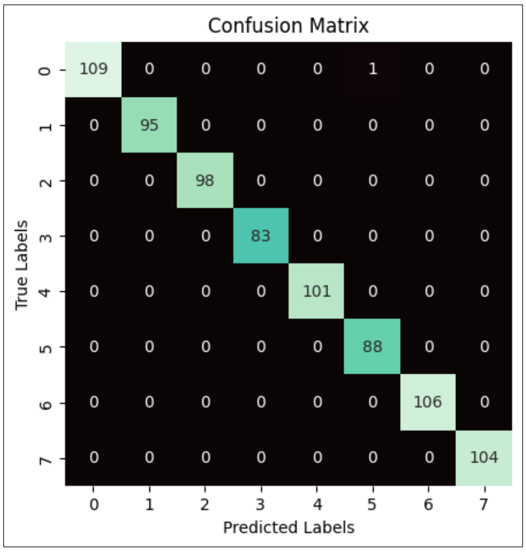


Figure 3-1 Cut flower Advisor confusion matrix

c. Classification report

The classification report offers summary metrics for the entire dataset along with metrics for each class, including precision, recall, F1-score, and support. The model is effective, as evidenced by the high precision, recall, and F1 scores for each class in this classification report (Figure 3-2).

	precision	recall	f1-score	support
Alstroemeria	1.00	0.99	1.00	110
Anthurium	1.00	1.00	1.00	95
Carnation	1.00	1.00	1.00	98
Chrysanthemum	1.00	1.00	1.00	83
Gerbera	1.00	1.00	1.00	101
Lily	0.99	1.00	0.99	88
Orchid	1.00	1.00	1.00	106
Rose	1.00	1.00	1.00	104
accuracy			1.00	785
macro avg	1.00	1.00	1.00	785
weighted avg	1.00	1.00	1.00	785

Figure 3-2 Cut flower Advisor classification report

d. API testing

The Flask API was tested using the Postman API platform, as in Figure 3-3.

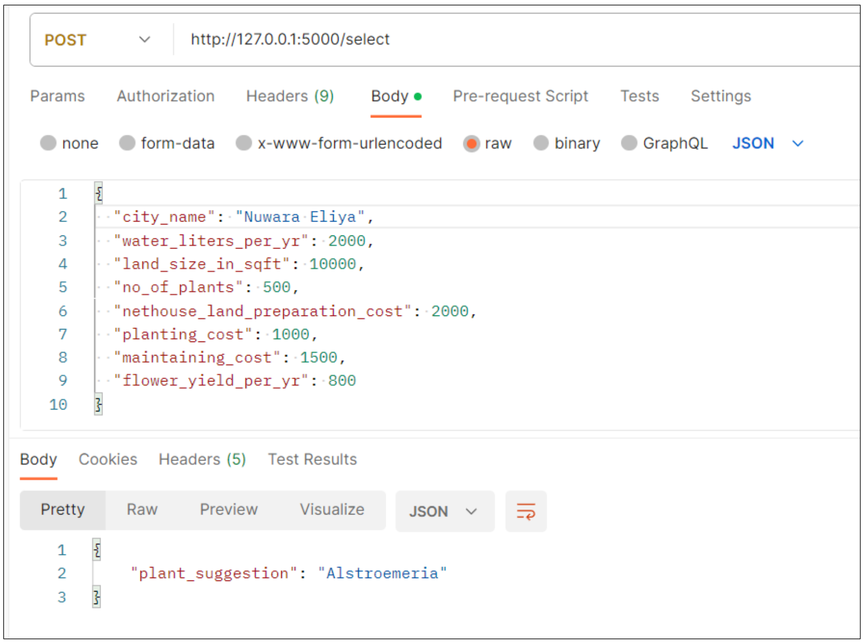


Figure 3-3 Cut flower Advisor API

- **Philo Variety Finder**

The Philodendron classification model was tested using 76 unseen test images. The binary classification model was tested using 152 unseen test images.

a. Model evaluation

The model object's ‘evaluate’ method is used to assess how well it performs given a certain dataset. Based on its learning parameters, the model uses the features in test images to produce predictions. To determine accuracy and loss, the predicted labels are compared with the actual labels of the images. The Philodendron classification model and the binary classification model exhibit good accuracy and minimal loss. Table 3-1 illustrates the accuracies and loss values of both models.

Table 3-1 Classification model evaluation results

Classification model	Accuracy	Loss
Philodendron classification model	0.8289	0.5133
Binary classification model	1.0000	0.1938

b. Confusion matrix

In order to calculate the confusion matrix, the scikit-learn confusion_matrix function is used to compare the actual class labels to the predicted class labels. A confusion matrix is generated by placing predicted class names on the x-axis and true class names on the y-axis. It displayed that the model is doing excellently in correctly identifying the majority of the samples. Figure 3-4 illustrates the confusion matrix for Philodendron classification model, and Figure 3-5 illustrates the confusion matrix for binary classification model.

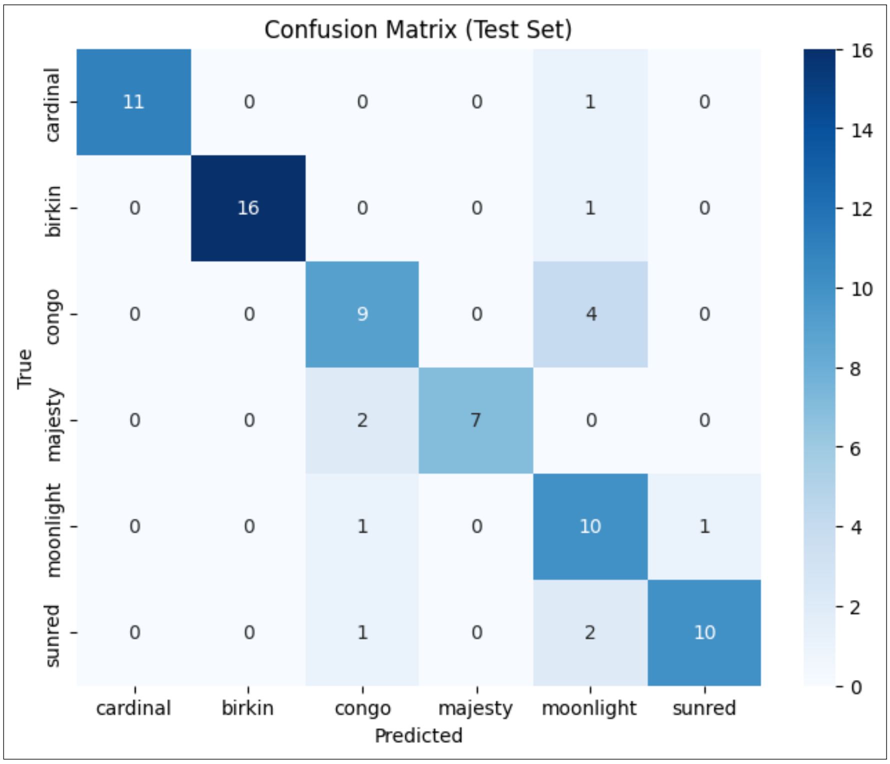


Figure 3-4 Philodendron classification confusion matrix

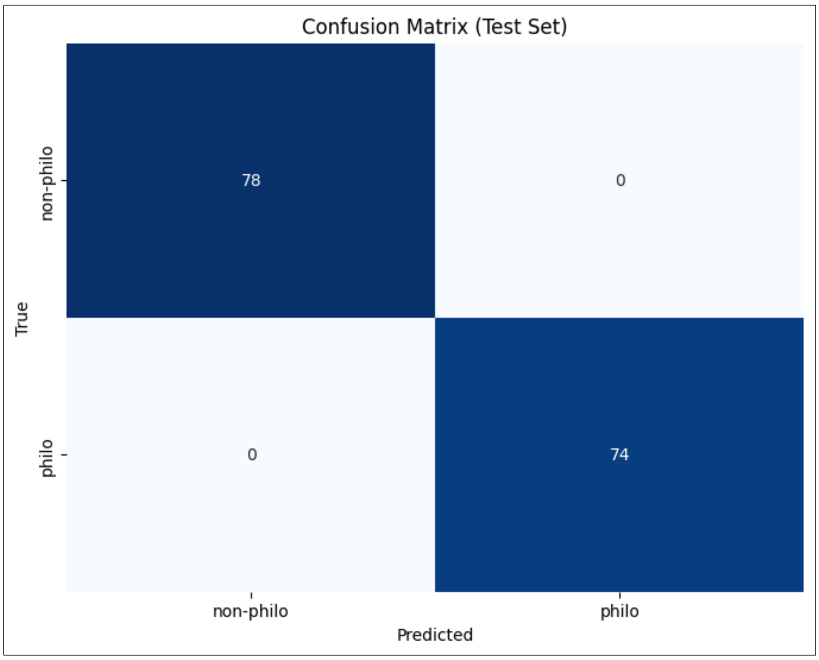


Figure 3-5 Binary classification confusion matrix

c. Classification report

Scikit-learn's classification_report function is used to generate the classification report. It provides metrics for each class, such as precision, recall, F1-score, and support. These measures are helpful for assessing the model's performance in each class. It is useful to know which classes the model does better in, and which ones might require some improvement. In Philodendron classification model, for certain classes, high precision and recall demonstrate the ability to correct predictions. The overall accuracy rate is 83% as in the classification report shown in Figure 3-6.

	precision	recall	f1-score	support
cardinal	1.00	0.92	0.96	12
birkin	1.00	0.94	0.97	17
congo	0.69	0.69	0.69	13
majesty	1.00	0.78	0.88	9
moonlight	0.56	0.83	0.67	12
sunred	0.91	0.77	0.83	13
accuracy			0.83	76
macro avg	0.86	0.82	0.83	76
weighted avg	0.86	0.83	0.84	76

Figure 3-6 Philodendron classification model classification report

In binary classification model, the model performs significantly in terms of classification, obtaining an excellent accuracy rating for the test dataset as displayed in Figure 3-7.

	precision	recall	f1-score	support
non-philo	1.00	1.00	1.00	78
philo	1.00	1.00	1.00	74
accuracy			1.00	152
macro avg	1.00	1.00	1.00	152
weighted avg	1.00	1.00	1.00	152

Figure 3-7 Binary classification model classification report

d. Sample testing of set of images

A sample of the test images is displayed along with the actual classes, predicted classes, and associated probabilities. It's a helpful technique for understanding the model's behavior on actual data. Figure 3-8 shows the results for the sample testing set of the Philodendron classification model. Figure 3-9 shows the results for the sample testing set of the binary classification model.

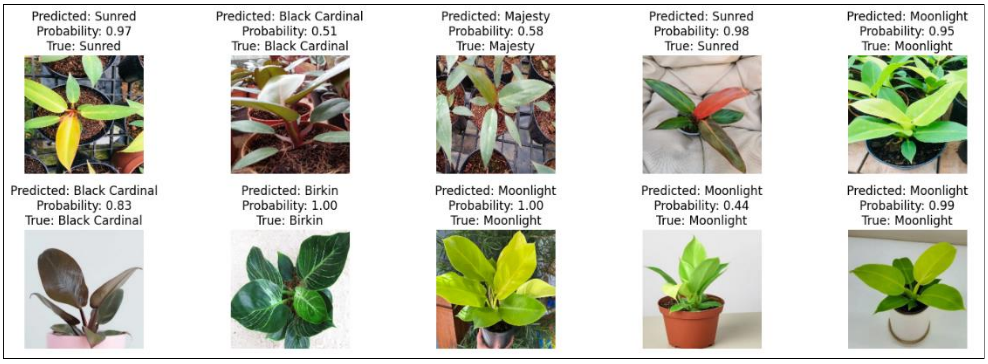


Figure 3-8 Philodendron classification model sample testing

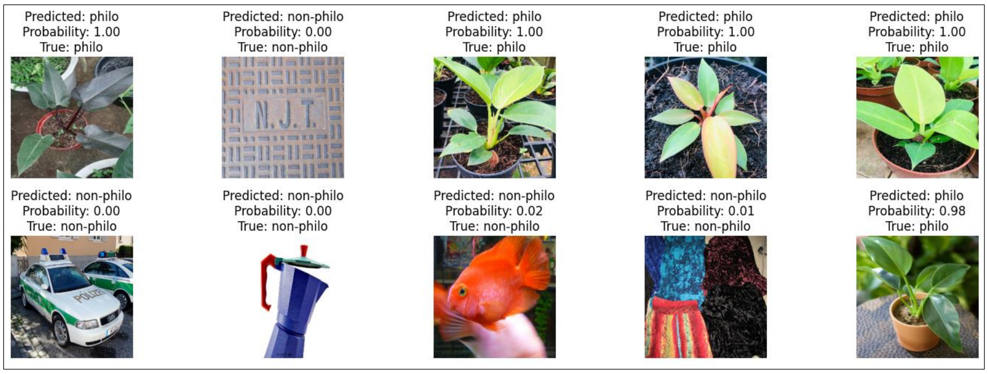


Figure 3-9 Binary classification model sample testing

e. API testing

The Flask API was tested using the Postman API platform. An image of a Philodendron variety was tested, and the output is illustrated in Figure 3-10. Result for an unknown image is shown in Figure 3-11.

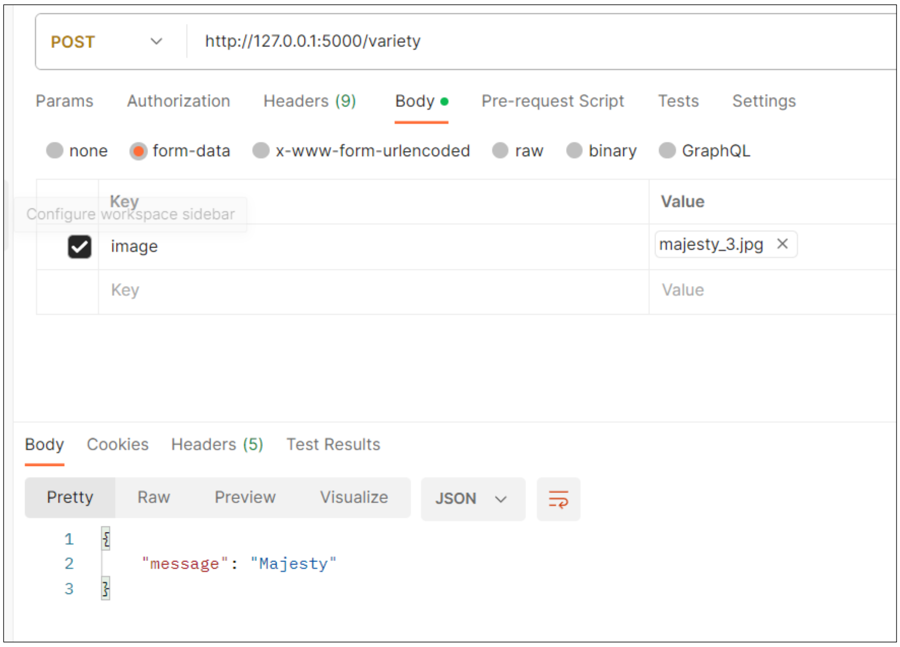


Figure 3-10 Flask API testing for philodendron

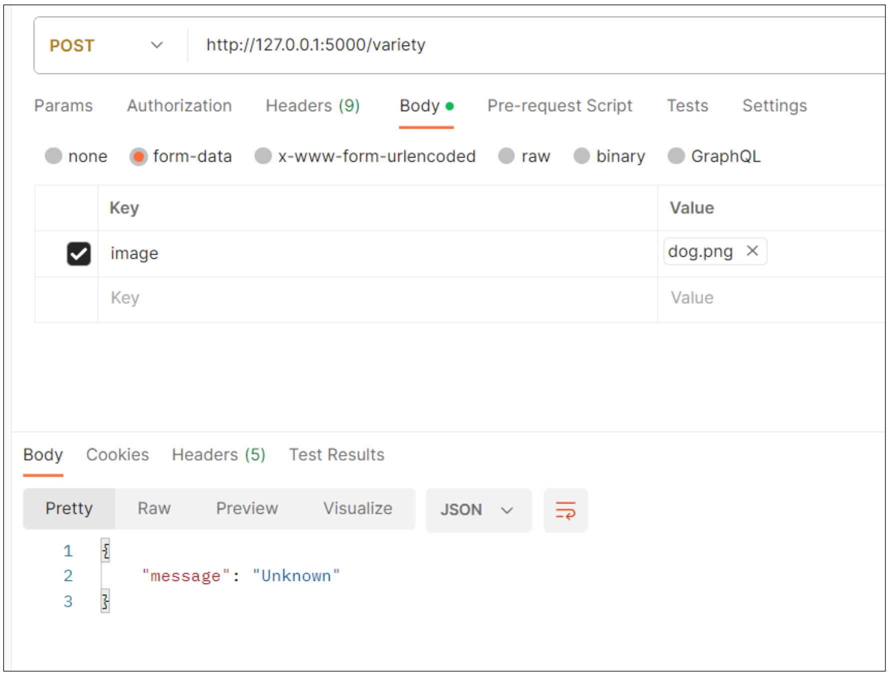


Figure 3-11 Flask API testing for unknown image

3.2 Research Findings

Plant Pal, a smartphone application created expressly to fulfill the needs of the floriculture industry with a particular focus on beginners as the individual component, is the outcome of the research. Two essential items, the Cut flower Advisor and the Philo Variety Finder, have been methodically developed as part of the full ‘For Beginners’ component.

The Cut flower Advisor uses machine learning, which is evidence of its accuracy and efficiency. This component expertly suggests the ideal cut flower crop from a variety of eight different types while carefully considering important inputs including the environment, resources available, and production goals. This level of accuracy enables growers to make well-informed decisions, which eventually improves crop cultivation results.

In addition, the Philo Variety Finder makes use of CNN and transfer learning, cutting-edge image recognition techniques. Its potential is demonstrated by its remarkable accuracy in identifying Philodendron plant varieties from images. Additionally, the addition of smart validation expands its functionality and ensures that customers get accurate information about the identified variety.

These findings highlight how effective Plant Pal is as a strong and smart system. Modern technologies could support the floriculture sector, as demonstrated by the Cut flower Advisor, which is powered by machine learning, and the Philo Variety Finder, which is powered by deep learning. Plant Pal is positioned to become an essential tool for both novice growers and enthusiasts, bridging the gap between inexperience and floricultural expertise by offering accurate recommendations and improving the learning aspect of plant variety identification.

3.3 Discussion

Two essential components, the Cut flower Advisor and the Philo Variety Finder, were carefully developed as part of the overall For Beginners module, demonstrating significant achievements in terms of advanced technology and model accuracy.

The Cut flower Advisor is supported by machine learning and achieves an impressive 99.87% accuracy on unobserved data. It serves as evidence of the app's technical skill, effectively selecting the best cut flower crop out of eight options depending on inputs including environmental factors, resource availability, and production goals. By enabling growers to make informed decisions, this level of accuracy significantly enhances crop production outcomes. Possibilities for the future include increasing the number of suggested crops and incorporating real-time environmental monitoring for improved accuracy.

The Philo Variety Finder uses CNN and transfer learning in parallel to obtain an 82.89% accuracy rate on unseen data, demonstrating its competence in image identification. The addition of smart validation, which guarantees users obtain accurate and trustworthy information about the recognized Philodendron plant variety, further strengthens its potential. Expanding the coverage of plant varieties and investigating deep learning’s use in broader plant identification scenarios could be potential future goals.

4 SUMMARY OF EACH STUDENT'S CONTRIBUTION

Table 4-1 Individual contribution

Student ID	Student Name	Contribution
IT19169736	Gamage M.G.U.D.	• Obtain the climate data without sensors based on the grower's location • Predict the best cut flower plant to grow, considering weather and infrastructure • Differentiate the main six varieties of Philodendron using an image of the plant. • Validate the user's input to determine if the plant falls within any of the six Philodendron varieties. • API development of the individual component • Cross-platform mobile application development of the individual component • Participating in UI designing

5 CONCLUSION

The floriculture industry stands as an important pillar of the Sri Lankan economy. While relying on manual methods and expertise to monitor growth, maintain export standards, anticipate supply and demand, select crops for cultivation, and identify plant varieties, growers within this industry still face numerous difficulties. These difficulties, which become even greater for newcomers to the floriculture industry, demand the existence of a smart system to accelerate operations and reduce possible losses. The Plant Pal smartphone application, a complete tool created to help change the floriculture environment, was developed by the team in response to these industry-wide needs. The Cut flower Advisor and the Philo Variety Finder are included in the application's specific For Beginners feature. These components show an effort to use cutting-edge technology and generate remarkably accurate models. Since cut flowers are popular among beginners, the cut flower adviser concentrated on choosing the ideal cut flower crop to grow. The focus of Philo Variety Finder is on newcomers' difficulties recognizing Philodendron varieties, which are extremely popular in the cut foliage industry.

The Cut flower Advisor uses machine learning to produce an impressive 99.87% accuracy rate on unobserved data. This accomplishment highlights the app's technical proficiency and capacity to suggest the best cut flower crop out of eight options depending on critical inputs, including environmental conditions, resource availability, and business objectives. This level of accuracy promises to considerably improve crop cultivation outcomes by enabling growers to make informed decisions. Future efforts might include broadening the selection of suggested crops and including real-time environmental monitoring. In parallel, the Philo Variety Finder employs deep learning, achieving an impressive 82.89% accuracy rate on unseen data in the domain of image identification. Its capabilities are strengthened by the addition of smart validation, ensuring users receive accurate and reliable information about identified Philodendron plant varieties. The plant variety dataset might be expanded, and CNN applications in more complex plant identification scenarios could be studied. Future improvements have significant potential. In conclusion, the development of the Plant Pal mobile application marks an important step in supporting the Sri Lankan floriculture industry. Plant Pal establishes itself as a key tool, bridging the gap between novice growers and floriculture experts by solving the various difficulties faced by both experienced and beginning growers. With cutting-edge technology, outstanding model accuracy, and a dedication to industry advancement, Plant Pal promises to bring in a new era of prosperity and achievement for the floriculture sector in Sri Lanka and beyond.

REFERENCES

[1] S. M. P. C. a. K. T. D. Padmini, "Present status and future scope of floriculture industry in Sri Lanka and its potential in women empowerment," *Sri Lanka Journal of Social Sciences,* pp. 31-40, 2017.

[2] S. B. A. G. D. P. A. Y. Prof. Swati Dhabarde, "Agricultural Crop Recommendation System Using IoT and M.L.," *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT),* vol. 2, no. 1, 2022.

[3] M. B. J. S. S. P. S. D. K. Sameer. M. Patel, "Smart Agriculture using IoT and Machine Learning," *International Research Journal of Engineering and Technology (IRJET),* vol. 8, no. 4, 2021.

[4] M. K. Sharma, S. Agrahari, S. Tyagi and S. Punia, "Crop Selection Using Data Analytics," in *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, Greater Noida, India, 2020.

[5] S. Jain and D. Ramesh, "Machine Learning convergence for weather based crop selection," in *2020 IEEE International Students' Conference on Electrical,Electronics and Computer Science (SCEECS)*, Bhopal, India, 2020.

[6] S. a. B. G. a. S. R. N. V. a. A. P. K. Dharwadkar, "Floriculture classification using simple neural network and deep learning," in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, India, 2017.

[7] R. Rathnayake, M. T. Pramodi, I. Gayathree, L. Rashmika, M. Gamage and A. Gamage, "BlossomSnap: A Single Platform for all Anthurium Planters Based on The Sri Lankan Market," in *2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, 2022.

[8] L. P. E. P. R. M. J. J. S. a. A. C. G. A. Carneiro, "Grapevine Varieties Identification Using Vision Transformers," in *IGARSS 2022 - 2022 IEEE*

International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 2022.

[9] A. F. K. a. C. Y. Y. Altuntaş, "Identification of Apricot Varieties Using Leaf Characteristics and KNN Classifier," in *International Artificial Intelligence and Data Processing Symposium (IDAP)*, Malatya, Turkey, 2019.

[10] E. A. D. E. S. L. Mr. Gayan Weerasinghe - Export Promotion Officer, "INDUSTRY CAPABILITY REPORT - Sri Lankan Floriculture Sector," Export Development Board (EDB), Sri Lanka, 2021.

[11] K. L. Getter, "Average Daily Temperature Effects on Paclobutrazol Efficacy of Four Floriculture Crops," *HortTechnology*, vol. 25, no. 6, p. 742–746, 2015.

[12] L. M. Mortensen, "Effects of air humidity on growth, flowering, keeping quality and water relations of four short-day greenhouse species," *Scientia Horticulturae*, vol. 86, no. 4, pp. 299-310, 2000.

[13] A. J. H. T. S. M. A. G. Z. K. B. C. U. Hemant Upadhyay, "Exploration of Crucial Factors Involved in Plants Development Using the Fuzzy AHP Method," *Mathematical Problems in Engineering*, p. 9, 2022.

[14] "SPACE NEEDED FOR A PLANT NURSERY," ACS Distance Education, 10 June 2019. [Online]. Available: https://www.studyacs.com/blog-space-needed-for-a-plant-nursery-1061.aspx.

[15] T. E. o. E. Britannica, ""philodendron"," Encyclopedia Britannica, 15 June 2023. [Online]. Available: https://www.britannica.com/plant/Philodendron. [Accessed 10 September 2023].

[16] S. C. S. M. L. S. C. Loss-Oliveira L, "Evolution of Philodendron (Araceae) species in Neotropical biomes," *PeerJ*, 2016.

[17] L. H. A.-B. N. Azizi MMF, "Integration of advanced technologies for plant variety and cultivar identification," *Journal of Bio Sciences*, pp. 46-91, 2021.

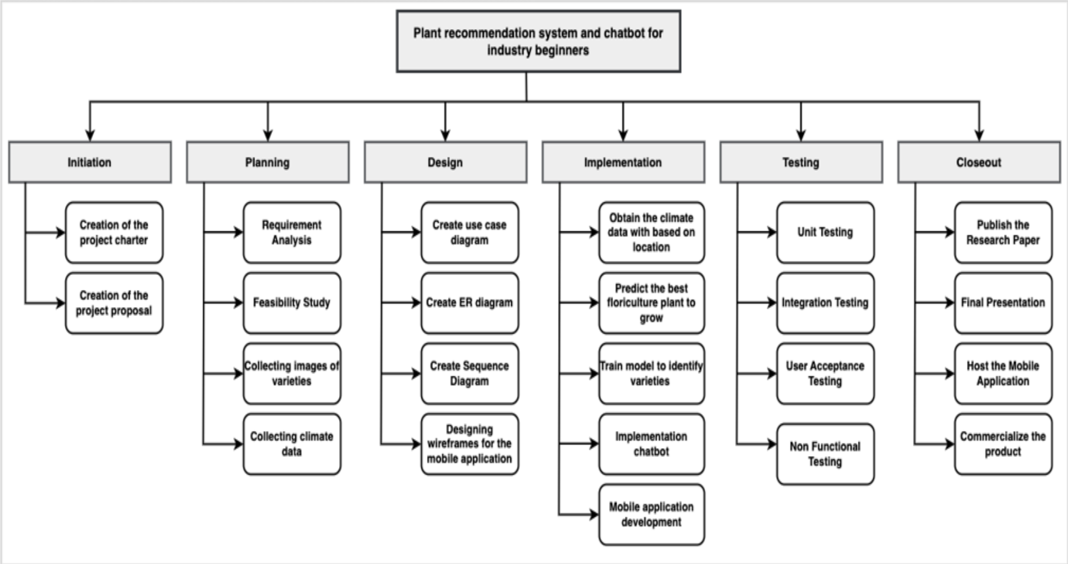
[18] "Sri Lanka Export Development Board," [Online]. Available: https://www.srilankabusiness.com/floriculture/cut-flowers.html.

APPENDICES

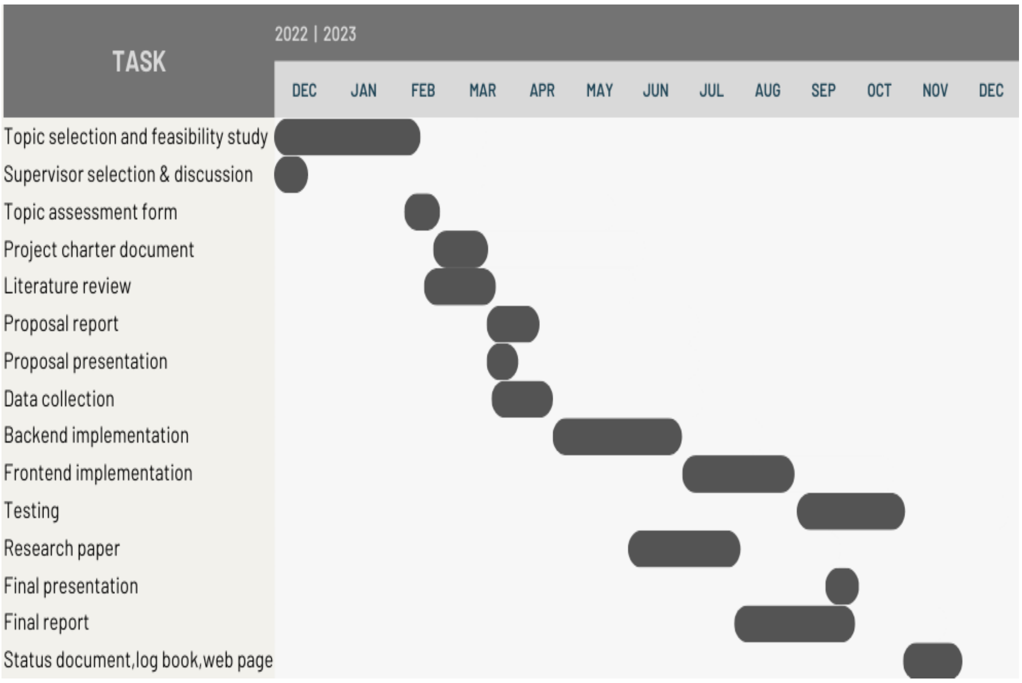
Appendix 1 : Survey on Floriculture Industry

https://forms.gle/mknPpztYp63e2wJE8

Appendix 2 : Individual Work Breakdown Structure



Appendix 3 : Gantt Chart



Appendix 4 : Budget Estimation

Component	Price
Travelling cost and other expenses	Rs. 30000
Cost for hosting the mobile app on Play store	Rs. 8075
Cost for hosting the mobile app on App store	Rs. 22610/monthly