SMART ASSISTANCE FOR THE FLORICULTURE INDUSTRY

Final Group Report

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 Group Report

Liyanage S.R.	(IT20005726)
Basnayake N.S.N.	(IT19994406)
Gamage M.G.U.D.	(IT19169736)
Prabhashi P.A.N.	(IT20017088)

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

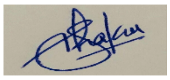
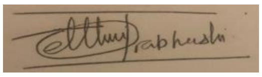
Sri Lanka

September 2023

DECLARATION

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

Name	Student ID	Signature
Liyanage S.R.	IT20005726	
Basnayake N.S.N.	IT19994406	
Gamage M.G.U.D.	IT19169736	
Prabhashi P.A.N.	IT20017088	

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

Signature of the Supervisor: Date:

……………………………. ..……………………….

Dr. Sanvitha Kasthuriarachchi

Abstract

The floriculture industry in Sri Lanka is a thriving economic venture, yet it is riddled with inefficiencies stemming from outdated technologies and manual methodologies. Particularly for newcomers, challenges abound in areas like plant variety selection, growth monitoring, adherence to export criteria, and precise plant identification. This research addresses these notable gaps, proposing a technologically advanced solution tailored to meet the sector's unique needs. Leveraging state-of-the-art technologies, including machine learning, deep learning, and computer vision, we have conceptualized a mobile application designed to simplify and optimize floriculture practices. This application boasts functionalities such as real-time plant growth prediction, in-depth leaf affliction analysis, informed plant recommendations, and accurate supply-demand forecasting. The rigorous system and user requirements detailed in this study underscore its commitment to precision, efficiency, adaptability, and user-friendliness. By integrating platforms and libraries like Keras, Scikit-Learn, OpenCV, Flask, and React Native, the system not only ensures robust technological backing but also guarantees a seamless user experience. In essence, this research, through its proposed mobile application, aims to revolutionize Sri Lanka's floriculture sector. By transitioning from traditional methods to a data-driven approach, it aspires to elevate industry standards, reduce barriers for entry, enhance profitability, and ultimately pave the way for a more sustainable and prosperous floriculture future in the country.

Keywords: floriculture, plant selection, beginners, variety identification, philodendron, machine learning, deep learning, computer vision, Sri Lanka, mobile application, export, leaf affliction, demand, growth monitoring

ACKNOWLEDGEMENT

We seize this moment to convey our sincere gratitude to all those who played a pivotal role in supporting and assisting us in the successful completion of the project.

First and foremost, our heartfelt appreciation goes out to the CDAP team and the esteemed research panel for their unwavering diligence and unwavering commitment throughout the course. Their dedication to ensuring a valuable educational experience and facilitating the successful culmination of our research project has proven to be indispensable. We extend special thanks to our project supervisors, Dr. Sanvitha Kasthuriarachchi, and co-supervisor, Dr. Lakmini Abeywardhana, for their steadfast leadership and invaluable guidance, which steered our research in the right direction.

We also extend our deepest thanks to our external supervisor, Mr. Anandathissa Illeperuma, for his invaluable advice and support throughout the project. Our gratitude extends to the entire team at Omega Green Pvt. Ltd. for their invaluable contributions in data collection, knowledge sharing, and priceless guidance. We wholeheartedly appreciate the ongoing support and collaboration of our colleagues in the university throughout this project.

Lastly, we wish to express our profound gratitude for the unwavering support and encouragement we received from our families and friends. Their belief in us has been a constant source of motivation, and we attribute our success in completing this project to their unwavering faith and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
ANN	Artificial Neural Network
API	Application Programming Interface
CNN	Convolutional Neural Network

DL	Deep Learning
GLCM	Gray-Level Co-Occurrence Matrix
HSV	Hue, Saturation, Value
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
ROI	Regions of Interest
SVM	Support Vector Machine
LightGBM	Light Gradient-Boosting Machine
CatBoost	Categorical Boosting
R^2	R-Squared
RMSE	Root Mean Square Error
MSE	Mean Square Error
XGBoost	Extreme Gradient Boosting

LIST OF APPENDICES

1 INTRODUCTION

Sri Lanka, often described as a jewel of nature's creation, stands out not only for its captivating landscapes but also for its burgeoning floriculture sector. This sector represents more than just an agricultural venture; it symbolizes the country's potential to shape socio-economic development, harness its unique biodiversity, and tap into a global market hungry for high-quality floriculture products. One of the country's most remarkable features is its diverse climatic regions within a small geographical span. From tropical rainforests to temperate highlands, such variations provide an unparalleled advantage for cultivating a wide variety of plants. The nation's fertile soils, enriched by its numerous rivers and water sources, further accentuate this advantage. These natural benefits, combined with human efforts and technological interventions, have positioned the Sri Lankan floriculture industry as a formidable competitor on the global stage.

Tracing back to 1980-1981, the country embarked on an export-oriented journey. Initial steps were modest, but with dedication and strategy, Sri Lanka began to establish its footprint in international markets. When it comes to the export market of the floriculture industry, Sri Lanka mainly aims at the European market, and as a percentage, it is 60%. Recently, other buyers have shown interest in Sri Lankan floriculture products, such as those from the Middle East, Japan, the USA, and Korea. [1]As the country healed from its internal conflicts and as the tourism industry grew, so did the domestic demand for floriculture products. From hotels to homes, the allure of vibrant flowers and ornamental plants became ubiquitous. The product range also diversified over time, spanning from Cut Flowers to Tissue Cultured Plants, each finding its niche in both local and global markets.

The floriculture sector is considered a lucrative agricultural venture in Sri Lanka, with the potential to drive social and economic development. However, many growers face challenges due to the lack of advanced solutions and reliance on outdated technology [2]. Critical aspects such as monitoring plant growth, upholding export standards, forecasting demand, crop selection, and plant variety identification currently rely heavily on manual processes and historical knowledge.

In the intensely competitive floriculture industry, both at local and international levels, the quest to deliver unparalleled quality while optimizing production costs is of utmost importance. At the heart of this endeavor are ornamental plants, the foundation for products such as cut foliage, flowers, and landscaping plants. Rigorous and regular monitoring of these ornamental plants is vital to ensuring they meet the exacting industry standards, all while streamlining human effort and associated costs. Metrics like leaf count, leaf color, and plant height are pivotal indicators of a plant's current growth phase. For an industry thriving on competition, securing both local and international commercial orders requires clear communication on delivery timelines. This implies that companies must comprehend the precise time a specific plant batch will mature to the required order standards. The ability to predict when plants will reach the desired supply stage is crucial for aligning with order schedules. However, the technological framework aiding growth monitoring in Sri Lanka's floriculture sector lags behind its international peers. The integration of a robust, IT-centric growth monitoring and prediction system stands as a promising solution, with the potential to significantly enhance the industry's performance and standing.

Plant cultivation is susceptible to issues like pest infestations and nutrient deficiencies that can significantly impact yields. Therefore, early detection of affected plants is essential for timely intervention through pest control and fertilization. Compliance with export standards, which often involve assessing the percentage of affected plants, is crucial to prevent financial losses from rejected shipments. Currently, these tasks are performed manually, and there is an absence of efficient systems to address them.

Predicting demand for ornamental plants presents its own set of challenges. Accurate demand forecasting is essential for the success of Sri Lanka's floriculture industry. Inaccurate predictions can lead to lost sales, excess inventory, and resource wastage due to overproduction. On the flip side, failing to meet demand can harm the industry's reputation. Precise forecasting, enabled by modern technology like predictive analytics and mobile apps, helps businesses adapt to changing floral preferences globally. This empowers them to allocate resources efficiently, minimize waste, and enhance profitability. It also aids in meticulous production planning, inventory management, and logistics, leading to more stable supply chains and heightened customer satisfaction. Despite the critical role of forecasting, the floriculture sector remains underexplored in this area, offering ample opportunities for further research and innovation.

The floriculture industry poses numerous challenges for newcomers who may struggle to decide which plants to cultivate based on their specific circumstances. Natural factors like temperature, water availability, and humidity vary depending on the location, impacting plant growth. Additionally, a floriculture producer's performance depends on factors such as financial stability, available personnel, and space. Unfortunately, there is no comprehensive system in place to educate beginners on these matters. Furthermore, distinguishing between different varieties of floriculture plants, such as Philodendron, can be challenging, and there is a lack of effective resources for learning this knowledge. Consequently, the floriculture industry in Sri Lanka urgently requires a smart assistance system.

To address these issues, this research examined various studies on floriculture management, machine learning, deep learning, and computer vision. The goal is to create a mobile app that can assist beginners and interns in the floriculture sector. This app will facilitate the calculation of the percentage of affected plant leaves, ensure adherence to export standards, monitor plant growth, predict demand, assist in crop selection, and aid in plant variety identification without the need for extensive industry expertise. Implementing such a system has the potential to not only save time but also transform the floriculture industry by reducing financial losses and enhancing plant quality and yield.

The upcoming sections will delve into the background and literature review, outline the research problem and gap, specify the objectives, detail the methodology, and provide system diagrams as part of this research effort.

1.1 Background & Literature Survey

Sri Lanka's floriculture industry boasts a diverse array of exceptional products and stands as a global frontrunner in this sector. In 2021, Sri Lanka garnered an impressive $16 million from floriculture exports [3]. While the island sees a surge in both small and large-scale floriculture ventures, many growers grapple with challenges stemming from outdated technology and a scarcity of intelligent solutions. Key considerations in the floriculture domain encompass plant growth monitoring, adherence to export standards, demand prediction, plant selection, and variety identification. Regrettably, most of these tasks continue to be executed manually.

One pioneering study titled "Early season prediction of within-field crop yield variability by assimilating CubeSat data into a crop model" harnessed CubeSat-based imagery and LAI maps. The research team integrated these images into APSIM using a particle filter. The resulting system can forecast plant growth levels and the time required to attain them, circumventing the limitations posed by cloud cover in traditional satellite data. However, replicating such a system in the Sri Lankan context, with its expensive satellite-based image datasets, proves challenging, and the study does not delve into other critical factors like temperature, humidity, and fertilizer levels [4]. Monitoring plant growth encompasses various techniques, with a keen focus on aligning with industry demands. Notably, global markets impose particular specifications on growers. For instance, the plant Livistonia rotundifolia, which is in high demand, is often expected by customers to possess five distinct leaves and a height ranging between 25 to 30 cm. [5] Thus, cultivating plants to match market standards is pivotal.

Different metrics can be used to gauge plant growth, with specific traits pinpointing the desired growth stage. Key growth indicators for ornamental plants (OP) include stem height, stem diameter, leaf area, leaf length, leaf width, and leaf count. 4)Alongside these, it's imperative to account for environmental factors that profoundly influence plant growth. Elements such as temperature, water availability, light intensity, humidity, and nutrient levels play a decisive role in the development of ornamental plants.
Another research endeavor explores an automated system for monitoring and controlling ornamental plants, employing fuzzy logic techniques and an IoT-based approach. This system necessitates costly sensors and a highly technologically controlled environment to regulate temperature, air humidity, and soil moisture [6]. Additionally, high-tech smart greenhouse concepts have been proposed in various studies, encompassing essential growth monitoring factors such as temperature, humidity, and soil moisture [7], [8], [9], [10]. Although these studies cover a broader agricultural spectrum, they might not align perfectly with Sri Lanka's floriculture sector [11], [12].

A research initiative [13] in Sri Lanka for vegetable crops grown in Batticaloa focused on calculating the affected area in leaves to combat pests. RGB images were converted to HSV, and subsequent steps were taken to identify the affected regions and count the pests on specific leaves. A CNN model was developed for pest type classification. In another study, diseases in wheat crops were detected by capturing aerial images using drone cameras. BGR images were converted to the RGB format, and the necessary steps were implemented to identify the affected areas, providing location-specific results for growers to take timely preventive measures [14].

An IoT and ML-based agriculture system [15] for crop yield prediction employed meteorological data, encompassing variables such as temperature, humidity, and soil moisture, to predict high-yield crops for specific field areas. An IoT sensor system collected the meteorological data, while ML algorithms facilitated crop predictions. Another study used IoT to measure factors like humidity, pH, and rainfall, training ML algorithms on this dataset to offer precise forecasts and recommendations for thriving plant varieties [16]. Crop selection was further enhanced by the Crop scoring algorithm, which aided farmers in choosing the most suitable crops based on factors like rainfall, soil type, cropping month, and location. Seasonal weather predictions were tackled using RNNs in conjunction with soil factors and weather characteristics in a different study [17]. Meanwhile, a deep learning vision transformer architecture was deployed to distinguish various Grapevine varieties using leaf-centered RGB images, and feature extraction was employed to classify Apricot leaves in yet another research endeavor [18], [19].

The Smart Intelligent Floriculture Assistant Agent (SIFAA) system stands out as a comprehensive solution, melding expert knowledge with state-of-the-art techniques, including deep learning. SIFAA excels in plant disease diagnosis, treatment recommendations, customer product suggestions via Reinforcement Learning, and demand forecasting to motivate cultivators. It employs a mix of Linear Regression, advanced ensemble LightGBM Regressors, and feature engineering techniques. SIFAA's capabilities encompass plant development monitoring, pest identification, export standards maintenance, demand prediction, plant selection, and proper care [20]. Additionally, research delving into sugar production in Sri Lanka, employing ARIMA models and other ML methods for production forecasting, demonstrates the potential to enhance accuracy using SVM models [21]. Sri Lanka's vegetable demand forecasts were made more accurate through a range of technologies, offering valuable insights into future agricultural needs and societal demands [22], [23].

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 [3]. 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.
Selecting the right crops for floriculture newcomers poses a critical challenge, as it can greatly impact the success of their ventures. Factors such as environmental conditions, location-specific suitability, resource availability, and the identification of diverse plant varieties all come into play. Making informed decisions requires consideration of variables like temperature, humidity, and water, with regional climates influencing plant choices. Inexperienced growers may struggle to navigate these complexities, potentially leading to financial losses and resource wastage. Additionally, distinguishing between various plant varieties, especially in Sri Lanka's rich flora, can be daunting for novices due to the lack of accessible resources for identification. Incorrect plant identification can result in maintenance issues, customer dissatisfaction, increased pesticide use, and hindered learning opportunities, emphasizing the importance of accurate variety recognition for the industry's health and sustainability.

In the floriculture industry, despite regular fertilizer applications, plant infections can still occur, posing challenges for exporters who must adhere to stringent requirements to avoid rejection by importing countries. These requirements include providing precise information about plant afflictions, their previous occurrence, and current control measures, often with specific limits and area details. Visual inspections alone may not yield accurate measurements, and failing to meet export standards can result in substantial financial losses, encompassing missed profits and shipment-related expenses. Additionally, monitoring the progression of plant afflictions is crucial for evaluating fertilizer effectiveness, enabling growers to make informed decisions about continuing or changing their fertilizer strategies.

In the dynamic landscape of the floriculture sector, an emergent research challenge was identified: the development of an adept growth monitoring and prediction system for ornamental plants, suited to the distinct nuances of the industry. This study's primary objective was to devise an intelligent solution tailored to address two pivotal queries within a cohesive framework. Firstly, it sought to unravel how the growth trajectory of ornamental plants might be continuously tracked to accurately ascertain their current stage of growth. Secondly, the research aimed to identify methodologies that could be employed to predict the precise time duration required for these plants to mature to their desired growth size, optimizing them for order readiness. Delving deeper, the research bifurcated into two main avenues. The growth monitoring facet leaned on a multitude of environmental indicators and plant characteristics, encompassing attributes such as plant height, leaf count, ambient temperature, and humidity levels, aspiring to deliver a reliable prediction of the plant's current growth stage. Simultaneously, the time estimation facet homed in on predicting the interval within which an ornamental plant would reach its optimal growth stature, streamlining the ordering process.

The floriculture industry in Sri Lanka has witnessed inaccurate predictions can lead to missed sales opportunities, excessive inventory, and resource wastage, impacting both profitability and the environment. Implementing an advanced demand forecasting system powered by cutting-edge technology could revolutionize the industry. This technology-driven approach would enable precise alignment of production and inventory management with market demands, reducing waste and enhancing sustainability. A survey of industry professionals revealing that a substantial proportion relies heavily on personal experience for demand forecasting, underscoring the need for structured, data-centric methodologies. Respondents express a clear belief in the necessity of an advanced demand forecasting system, highlighting potential limitations in current forecasting methods. By embracing such a system, businesses can improve decision-making across production, pricing, and marketing, ultimately boosting sustainability, profitability, and competitiveness. Furthermore, there is openness among industry stakeholders to embrace advanced technologies like Artificial Intelligence (AI) and Machine Learning for demand forecasting. These technologies promise to usher in a data-driven era, uncovering intricate demand patterns, adapting to market dynamics, and enhancing decision-making processes. This transformation could significantly benefit the floriculture industry, reinforcing its competitive edge in an evolving market landscape while contributing to broader business growth and development.

1.3 Research Gap

Following is a description of related research conducted both locally and internationally, with the research gaps presented separately in the below subsections.

1.3.1 Leaf affliction analysis and shipment recommendation system

Research A:

The research [24] is related to identifying pests. It was conducted for the black orchid species in Indonesia. After uploading an image, the type of pest that attacked the leaf will be identified, such as ladybugs, mites, snails, or caterpillars. This was achieved using the Naïve Bayes algorithm for image processing, and the GLCM method was used for leaf texture feature extraction.

Research B:

The research [25] is related to detecting plant diseases in leaves using machine learning. For this purpose, they used a CNN model. After training the model on affected and healthy plants, the diseases were classified according to the models. However, this approach was only applicable to leaves and not to the plant body.

Research C:

The research [26] is related to detecting nutrient deficiencies in plants using machine learning. In India, this was done for the tomato crop using an ANN model. After examining the leaves, the correct nutrient deficiency was identified. The final output was a mobile app that could take and upload photos of the plant and provide information on the deficiency. It was observed that different fertilizers were used for different tomato plants, for instance, more nutrients were used during the flowering period.

Research D:

The research [27] is related to detecting nutrient deficiencies in corn leaves. The corn leaf image data was pre-processed, and RGB and HSV were used to finalize the data. Then, the data was trained with an SVM model and classified. Based on the classification, the appropriate nutrient fertilizer was recommended for the particular plant.

Research E:

The research [13] is related to calculating the affected area in a leaf. This was done in Sri Lanka for vegetable crops grown in Batticaloa to control pests after identifying the threat. First, the RGB images were converted to HSV. Then, the necessary steps were taken to detect the affected region, and based on that, the number of counts in that particular leaf was obtained. A classification model was created using a CNN model to identify the type of pest.

Research F:

The research [14] is related to detecting diseases in wheat crops. When there are large-scale crops, it is difficult to identify diseases by inspecting them one by one. Therefore, this research found a solution by obtaining aerial images using drone cameras. Using OpenCV, the BGR images were converted to the RGB format, and the necessary steps were taken to detect the affected area. After obtaining location-wise results, growers can take necessary actions to stop or prevent the spread of the disease.

Table 1-1 simply shows comparison between my system and the past research that I mentioned earlier.

Table 1-1 Leaf affliction analysis past research

Research Features	A	B	C	D	E	F	Proposed system
Finding the HSV ranges of the healthy leaf	-	-	-	-	-	-	√
Track affliction process	-	-	-	-	-	-	√
Calculated the affected area of a leaf in percentage	-	-	-	-	√	-	√
Recommendation system shipment	-	-	-	-	-	-	√
Mobile application	-	-	√	-	-	-	√

1.3.2 Cut flower selection and philodendron variety identification system

Research [28] 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 [29], 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 [17], 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.

[30] 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.

[31] 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.

[32] 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 [18], 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.

[19] 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.
Table 1-2 compares the proposed system with past studies.

Table 1-2 Crop selection and variety identification past research

Reference / Features	[28]	[29]	[17]	[30]	[31]	[32]	[18]	[19]	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	-	✓	-	-	-	✓	-	-	✓

1.3.3 Plant Growth Monitoring and Estimation of Growth Duration system

The floriculture industry has seen significant development around the world, largely due to technological advancements. However, the impact of the IT sector on the industry has been limited to certain areas. The proposed component aims to develop a smart mobile-based solution for growth monitoring and prediction of ornamental plants, taking into account the capacity and feasibility of the floriculture industry in Sri Lanka. Based on the background and literature survey, a research gap has been identified between the studies conducted thus far and the proposed component.

The first research identified is namely "Early season prediction of within-field crop yield variability by assimilating CubeSat data into a crop model". In that research, the research party used CubeSat-based images and LAI maps. The images were integrated into APSIM using a particle filter. The final system can forecast how long the plants will take to reach the desired level of growth, and the approach overcomes the impacts of cloud cover on traditional satellite data. But when it comes to the practical situation in the Sri Lankan sector, it is hard to achieve such a system with a high-cost satellite-based image data set, and the study does not consider other facts like temperature, humidity, fertilizer level, etc. [33]

The second research is about an automated system for controlling and monitoring ornamental plants using the fuzzy logic method. In that case, the proposed system is an IoT-based solution with actuators to control the temperature, air humidity, and moisture level of the soil. For that solution, expensive sensors and a highly technologically controlled environment are required. [34]The study "Information technology-controlled greenhouse: A system architecture" introduced a smart greenhouse equipped with actuators, utilizing NTP technology to control main factors including nutrients, soil moisture, temperature, and humidity. [35]

The next research is titled "Deep Learning for Image-Based Plant Growth Monitoring: A Review". In that case, a model using deep learning like CNN or RNN is used to predict the current growth state of the plant. But the research done on vegetable plants

like tomatoes, blueberries, etc. [36] The next research is about the manipulation of the rice L-galactose pathway. In that study, they have identified the relative growth rate of plant height, but the research aims at the paddy field and mainly considers the salt stress tolerance of the plant. [37]
Then there was research conducted to measure plant growth monitoring using wearable sensors. The system proposes two wearable sensors based on fiber Bragg gratings (FBGs), which will capture environmental factors such as temperature and humidity. [38]

The next research considered is titled "Controlling and monitoring ornamental plant care remotely using the Android application". In that research, they have proposed a highly technical IoT-based solution for controlling temperature, air and soil humidity, and water supply. But it only considered controlling the environmental factors. [39] There was research conducted to monitor the growth of orchid plants to predict the room humidity for the orchids. In that case, an IoT-based solution is proposed with sensors and actuators mainly to identify the suitable temperature and humidity. [40]

Table 1-3 Comparative Analysis of Previous Research on Growth Monitoring and Predictions

Application Reference	Identification of the current state of growth in ornamental plants	Consideration of the plant height	Identification of the current leaf color	Predicting the time required for plants to reach desired growth size for orders	Mobile app kind solution in Sri Lankan Sector
[33]	X	X	X	✓	X
[34]	X	X	X	X	X
[35]	X	X	X	X	X
[36]	✓	X	X	X	X
[37]	X	X	X	X	X
[38]	X	X	X	X	X
[39]	X	X	X	X	✓
[40]	X	✓	X	X	X
Proposed System	✓	✓	✓	✓	✓

1.3.4 Demand Prediction for ornamental plants

The following will show a certain gap between the proposed system and the currently used in already done research related to the area. Most of this research has been done in agriculture, not floriculture. Additionally, none of these will be discovered in Sri Lanka. Most of these technological advancements are novel to our country, as many of them were accomplished in India. According to SIFAA is a solution that can help the industry overcome these challenges. It is a sophisticated technology capable of diagnosing diseases affecting flowers, recommending treatments, and answering questions about floriculture. This system can efficiently provide reliable and intelligent information to farmers and can rapidly address complex issues. Additionally, it uses deep learning suits to aid ornamental plant disease identification through smartphones. To help customers find products that match their interests, a new recommendation system based on RL is proposed. It uses advanced demand forecast models to increase customer satisfaction and time wise fulfilment of user expectations. They have also done demand forecasting for the market, but they have used different kinds of technologies to implement it. [2]

This research is done in Sri Lanka, yet it is about sugar demand forecasting even though there is a slightly different in the technologies that we are going to use for forecasting the demand, but any other component will not be done in this research. [3]

This is also done in Sri Lanka, which has been done in the agricultural sector to predict both demand and supply for vegetables. [4]

In India, supply forecasting has been conducted by forecasting the harvests using a variety of technological approaches. [5]

All these were only found in agricultural products in the areas of demand and supply prediction category but didn't cover the other parts of this proposed system.
As for GPS technology that I am going to use in my system, it has never been done in the floriculture industry, but it has been done in the agriculture sector. [6]

This research is used to take the time series analysis for predicting the demand for the agriculture sector in India. [7]

In addition, this research is identical to the one that came before it; the only difference is that this one uses a mobile application and its primary focus is on ensuring customer satisfaction as the project's end objective while attempting to predict demand. [8]

The location of these research and investigations was India, and all are get researched on the agriculture sector, yet they have used different kinds of technologies like IOT, machine learning Decision trees, likewise. [9] [10] [11]

As for this last research, they also got the GIS technology, yet it is also used for the agriculture sector. [12].

In the larger landscape of research endeavors, it becomes evident that a considerable portion of scholarly work and technological advancements has primarily revolved around the agricultural sector, with a significant focus on India as a key player in this domain. This geographical concentration implies that many of the innovative technologies, methodologies, and concepts generated through these research efforts have the potential to be groundbreaking and uncharted territory when introduced in Sri Lanka.

What's especially noteworthy is the unique position of the floriculture industry within this context. As an industry that deals specifically with the cultivation and trade of ornamental plants and flowers, it stands apart from traditional agriculture in several ways. Therefore, the infusion of these technologies and ideas from the broader agricultural sector into the floriculture industry of Sri Lanka represents a truly pioneering and transformative endeavor.

This transition brings with it the excitement of exploring uncharted waters, where established practices may need adaptation or even a complete overhaul to align with the distinctive demands and intricacies of floriculture. It not only opens doors to innovation but also fosters a dynamic environment where new approaches and solutions can flourish, ultimately contributing to the growth and evolution of the floriculture industry in Sri Lanka.

Table 4 - Comparison between previous researches

Features / Reference	Demand forecasting according to different factors	Mobile Application
[41]	✓	✓
[42]	✓	-
[43]	✓	✓
[44]	-	-
[45]	-	✓
[46]	✓	-
[47]	✓	✓
[48]	✓	-
[49]	-	-
[50]	-	✓
[51]	-	-
Proposed System	✓	✓

1.4 Research Objectives

1.4.1 Main objective

Automated mobile application to provide smart assistance for the floriculture industry.

1.4.2 Specific objectives

- Growth monitoring and prediction system for ornamental plants.
- Leaf affliction analysis and shipment recommendation system.
- Plant recommendation system and variety identification for industry beginners.
- Demand prediction for ornamental plants.

1.5 Project Requirements

1.5.1 Functional requirements

- System should be able to get necessary data from the user.
- System should be able to capture a photo of a plant using the mobile phone camera.
- System should be able to select a photo from mobile phone gallery.
- System should be able to validate the user inputs.

1.5.2 Non-functional requirements

- The system must deliver precise classifications and offer dependable suggestions.
- The system needs to efficiently process images with speed.
- The system should demonstrate the capability to manage substantial data loads and adapt as required.
- The system should feature a user-friendly interface for easy utilization.
- The system should effectively address potential problems, including unsuccessful image uploads.

- The system should effortlessly oversee the mobile app's operations from any location.

1.5.3 System requirements

- Machine learning, deep learning, and computer vision development are conducted using Google Colab.
- Model development relies on TensorFlow, Keras, Scikit-Learn, and OpenCV libraries.
- Backend API development is carried out using PyCharm as the Integrated Development Environment (IDE) and Flask as the framework.
- API testing is performed using Postman.
- Front-end development is undertaken with Visual Studio Code as the IDE and React Native as the framework.
- App development is facilitated through the use of EXPO.

1.5.4 User requirements

- The user should be able to enter details to the app, for making predictions.
- The user should be able to capture a picture of a plant with their smartphone camera.
- The user should be able to upload images of relevant plants.

2 METHODOLOGY

To offer smart assistance to the floriculture industry, a smartphone application called 'Plant Pal' has been developed. This app includes functionalities such as tracking plant growth, ensuring export standards for afflicted plants, predicting demand, and providing guidance for newcomers. The general system architecture is depicted in Figure 2-1.

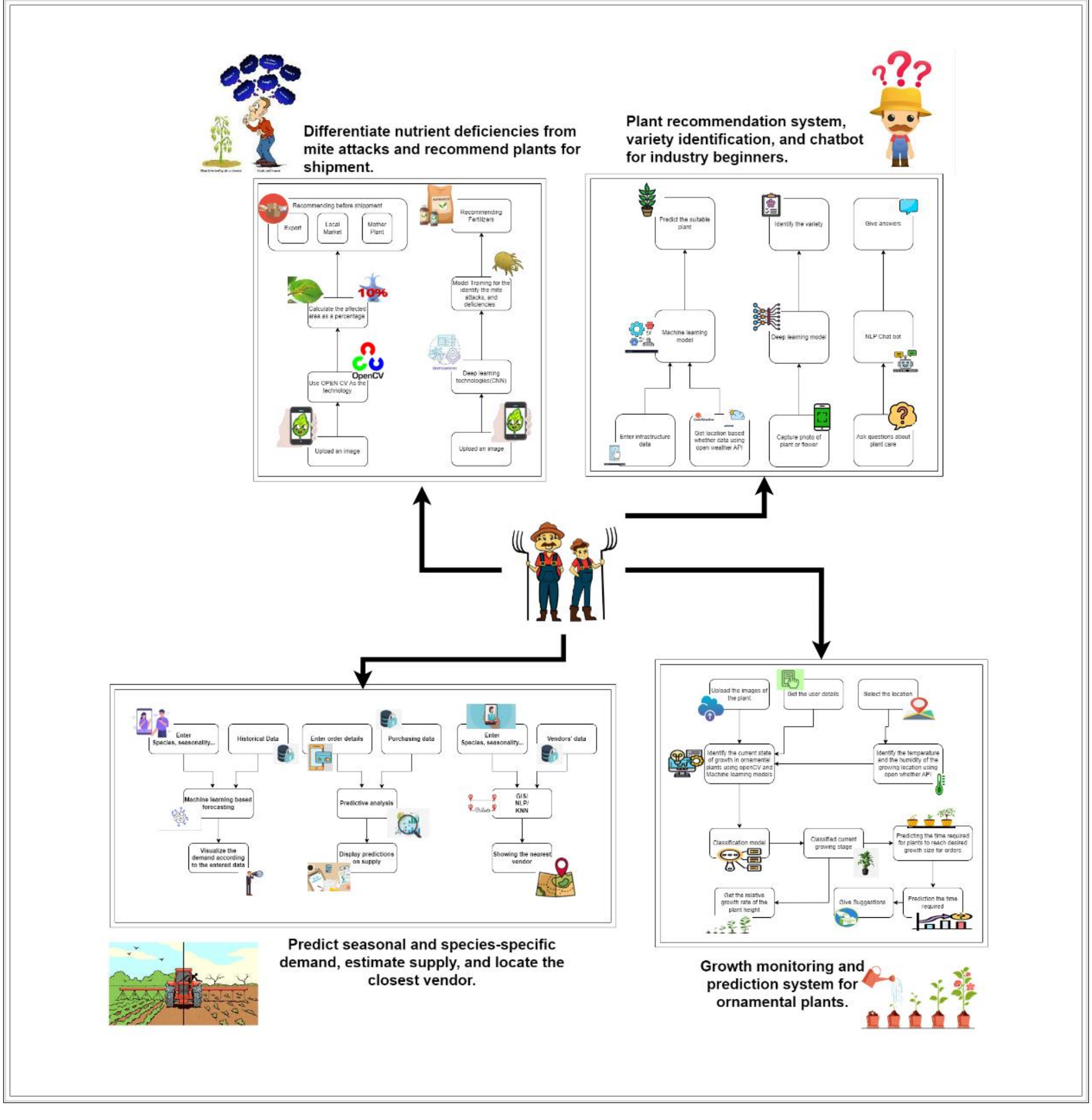


Figure 2-1 Overall system architecture

The system has been developed with the Agile software development process shown in Figure 2-2.

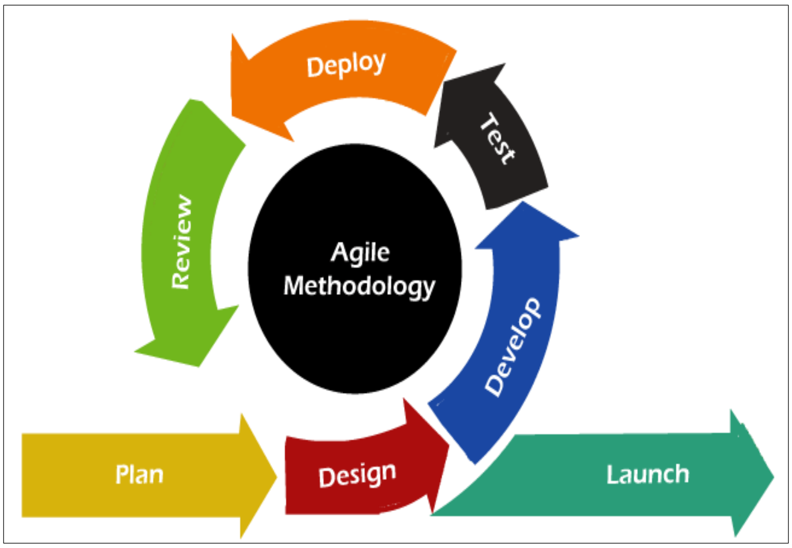


Figure 2-2 Agile methodology

The table 2-1 presents a comprehensive list of the primary technologies and corresponding tools/software utilized during the system implementation. These were carefully chosen to ensure the efficient development and optimal performance of the system. They range from programming languages and frameworks to specific software and mobile applications that aided in the creation, testing, and fine-tuning of the system.

Table 2-1 Utilized Technologies and Tools

Technologies	**Tools & Software**
React Native **Python** **JavaScript**	Visual Studio Code Jupiter Notebook Postman

Tensorflow **Matlab** **Flask Server** **Visual Studio Code** **Expo Framework** **Axios**	PyCharm IDE Expo Mobile App

2.1 Commercialization & Business Plan

Plant Pal, a smartphone app tailored for the floriculture sector, boasts four attractive features: leaf affliction analysis and shipping recommendations, cut flower selection and philodendron variety identification, growth monitoring, and supply-demand predictions, catering to both novices and ambitious entrepreneurs. It seamlessly operates on both iOS and Android platforms and doubles as a versatile backend API.

Plant Pal's business strategy encompasses subscription offerings, strategic collaborations, data monetization, and continuous development, with a primary focus on aiding interns and beginners in the floriculture domain and empowering budding entrepreneurs. Plant Pal aims to carve a niche in the market, positioning itself as an indispensable tool for industry enthusiasts while ensuring sustainable business growth. This objective is achieved through the provision of valuable insights and fostering user loyalty.

2.2 Backend Implementation

2.2.1 Leaf affliction analysis and shipment recommendation system

To calculate the affected area, images of plant leaves with various types of affections were collected. For the pursuit of precise leaf affliction detection, a multi-phased methodology was employed. Throughout these phases, a singular leaf image was subjected to a diverse array of processing and analytical techniques.

1) Color detection without background removal:

   The leaf image was initially transformed from its RGB format into the HSV format. This transformation, renowned for its capability to facilitate detailed color-based segmentation, enabled the demarcation of a specific HSV range. Such a range was delineated to encompass deep reds transitioning to slight oranges. A binary mask was consequently generated, in which the regions matching the stipulated HSV range were highlighted in white, while all other areas remained black. A visual representation is presented in Figure 2-3, placing the mask with the original image, thereby clarifying potential affliction areas.

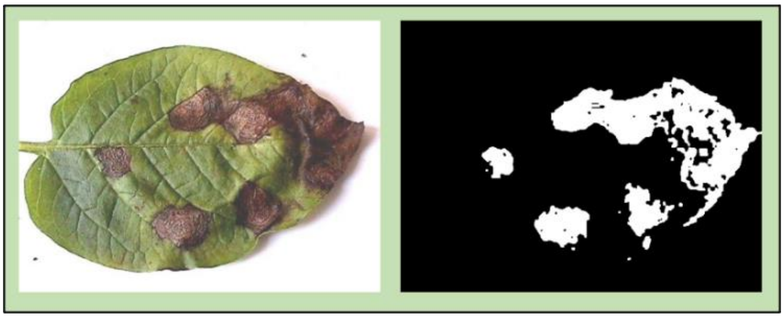


Figure 2-3 Original image and resultant mask

2) Color detection with background removal:

   To enhance clarity, the leaf was meticulously isolated from its background. The image was transmuted into grayscale, a process that simplifies its intrinsic structures. Following this transformation, a binary threshold was applied, yielding a binary representation. The contours within this binary construct were then meticulously detected, and the most prominent contour, assumed to be the leaf, was systematically isolated. The afore stated HSV range was reapplied to this isolated image, offering an enriched perspective of afflicted regions as shown in Figure 2-4.

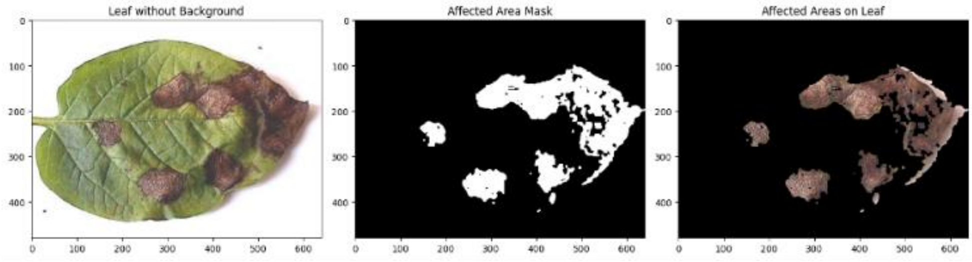


Figure 2-4 Resultant masks

3) Edge detection:

   A distinct phase dedicated to edge detection was initiated. The RGB image underwent a grayscale conversion, post which adaptive thresholding was employed, leading to the creation of a binary mask. The most dominant contour within this mask was identified, serving as a blueprint to craft a mask specific to the leaf. Subsequently, the Canny edge detection technique was deployed on the grayscale rendition of the isolated leaf, capitalizing on its proficiency to detect abrupt intensity variances. The resultant masks are shown in Figure 2-5.

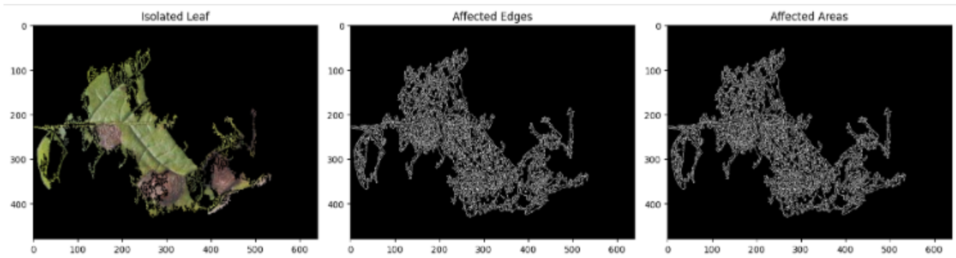


Figure 2-5 Edge detection resultant masks

4) Hybrid detection:

   A synergistic approach was adopted, amalgamating both color and edge detection paradigms. The image, once transmuted to grayscale and its contours emphasized, was processed using both the HSV range and the Canny edge detector. The outcomes of these dual techniques were amalgamated using a bitwise 'OR' operation, culminating in a composite mask that integrated insights from both detection strategies. The resultant masks are shown in Figure 2-6.

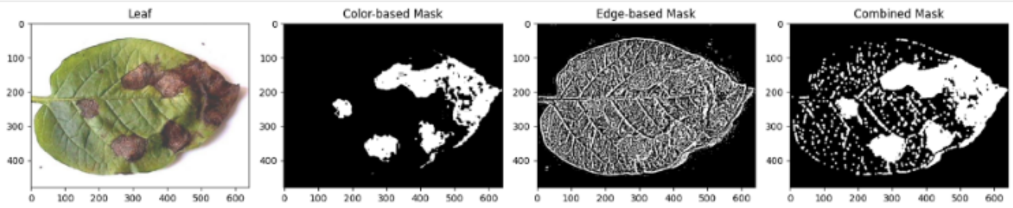


Figure 2-6 Hybrid detection resultant masks

5) Health-centric color segmentation:

i) Determination of HSV ranges for healthy leaves:

A curated collection of images, each representing diverse shades of healthy green, were collated. Specific Regions of Interest (ROIs) were assiduously selected, encapsulating characteristic leaf hues. Subsequent to these selections, an extensive evaluation was undertaken to determine the minimum and maximum values for each component of the HSV color spectrum. This rigorous analysis precipitated the establishment of an HSV range emblematic of a healthy leaf as shown in Figure 2-7.

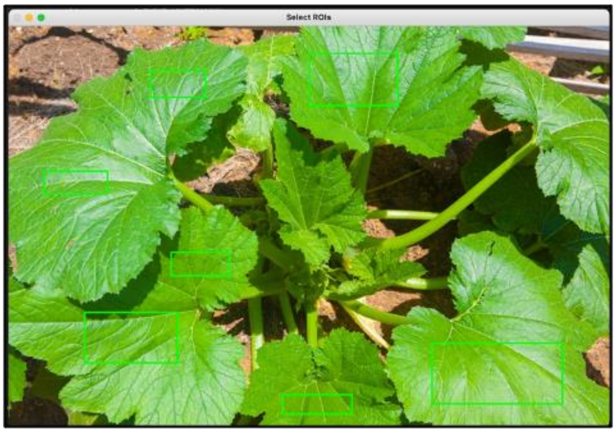


Figure 2-7 Selecting ROIs

ii) Color detection for healthy leaves:

The delineated HSV range, emblematic of a healthy leaf, was superimposed onto the isolated leaf image. This process demarcated the pristine regions of the leaf.

The inverse of this mask was then harnessed to highlight potential areas of affliction, thereby offering a granular overview of leaf health. Resultant masks are displayed in Figure 2-8.

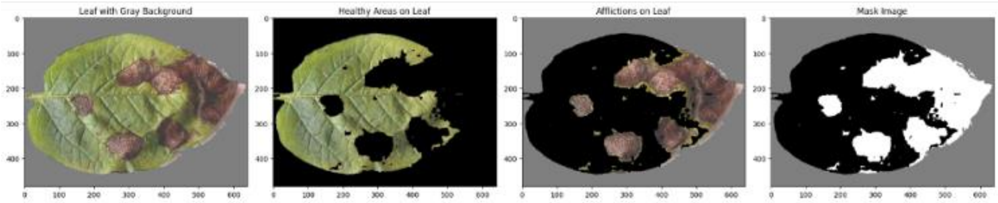


Figure 2-8 Resultant masks for healthy color detection

Each phase of this methodology was supplemented by illustrative visual aids in the guise of mask images. The intention was to transition from overarching detection methods to granulated, refined techniques, aspiring for the zenith of precision.

Upon examination of the derived mask images, it was discerned that the mask emanating from the healthy leaf HSV range phase manifested the most impeccable results. This optimized mask not only demonstrated adeptness in background obliteration but also accentuated the afflicted zones with unparalleled precision.

From the generated mask, an affliction ratio was deduced by tallying the white pixels (indicative of the afflicted regions) and juxtaposing them against the entire pixel count of the leaf. This ratio, when magnified a hundredfold, proffered the percentage of the leaf that bore afflictions. Such a computation leveraged the dichotomous nature of the mask.

*Affliction percentage = (Afflicted Leaf Area / Total Leaf Area) * 100*

2.2.2 Cut flower selection and philodendron variety identification system

1) Cut flower selection system

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.

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%. Feature selection was performed based on F-scores, which measure the difference in means between classes relative to the variance within each class. Then the top 9 features were selected for classification.

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. 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. The selected Gradient Boosting model was trained using the previously selected best 9 features, and hyperparameter tuning was performed as well. The improvement in model accuracy is visualized in this Figure 2-9.

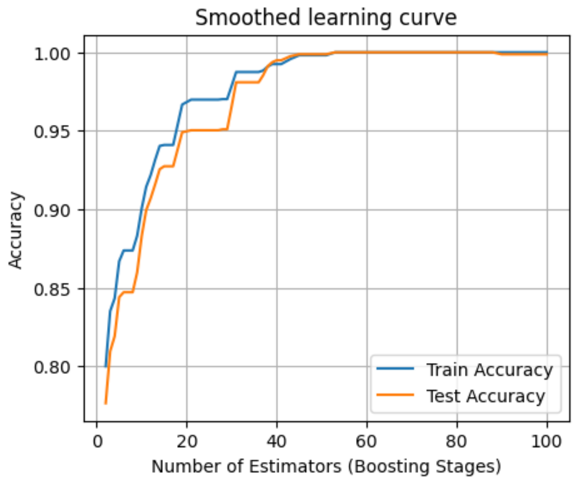


Figure 2-9 Learning curve for cut flower selection

OpenWeather API was used to fetch current weather data, including temperature (in Celsius) and relative humidity, for a given city.

2) Philodendron variety identification system

Mainly, two datasets were created: the Philodendron classification dataset and the Binary classification dataset. Philodendron classification 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. These images included random objects extracted from the ImageNet dataset and also images of Philodendron plants, which are out of the 6 varieties mentioned. Data were preprocessed with resizing and rescaling and were split into training, testing, and validation. Using the Keras API of TensorFlow, data augmentation techniques were used to 2 already-prepared image datasets.

Two CNN model architectures were built for Philodendron classification, a transfer learning model and a custom CNN model. A pre-trained MobileNetV2 model from TensorFlow Hub was used as the feature extractor. The 2 model architectures were separately trained using the prepared dataset of Philodendron images from 6 classes. An evaluation of the performance of the 2 trained models on the test dataset was done. The transfer learning model performed better, so the transfer learning model was selected for further development. Then the hyperparameter tuning was performed. Using these hyperparameters, the best model was built and trained again. And the learning curve displayed improved accuracy as in Figure 2-10.

Figure 2-10 Learning curve for philodendron 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 binary classification model was trained using the prepared binary classification dataset. Plots for training and validation accuracy as well as training and validation loss over the epochs are illustrated in Figure 2-11.

Figure 2-11 Learning curve for binary classification

2.2.3 Plant Growth Monitoring and Estimation of Growth Duration system

1) Growth Monitoring and Classifying System of Ornamental Plants

I. System Component Understanding

Embarking on this innovative journey, the Growth Monitoring and Classification System emerged as a cornerstone of the research. Designed with an aim to revolutionize growth monitoring within the Floriculture Industry (FI), the system was conceived to navigate the inherent complexities of this sector. As the FI demands significant investments in ensuring that plants are nurtured from their inception to their readiness for market orders, the intent was to introduce a system that could automate this meticulous monitoring process. Advanced technologies, notably machine learning, were employed to bring this objective to fruition.

II. Data Collation and Dataset Formation

A collaborative approach was adopted, with data being procured from the esteemed floriculture company, Omega Green (PVT) Ltd. The dataset's richness is attributed to its focus on five pivotal ornamental plants, each handpicked for their considerable demand both within local and global markets. Each plant underwent rigorous evaluation, with a series of attributes like plant type, variety, current height, characteristic leaves, current leaf color, and more serving as key determinants. The goal was to delve deep into the plant's growth trajectory and assess its 'current growth category', a classification attribute that encapsulates the current growth stage of the plant.

III. Rigorous Data Preprocessing and Cleaning

The preprocessing endeavor was seamlessly executed on the dynamic Google Colab platform, encompassing several integral stages:

IV. Data Comprehension

Prior to any analytical undertakings, it was imperative to understand the dataset's intrinsic structure. This step involved a meticulous assessment of potential correlations among various features, offering invaluable insights into their interdependencies.

- Null Value and Duplicate Rectification:

Data integrity was paramount. A rigorous protocol was followed to identify and subsequently rectify any null values and duplicate entries that could compromise the analysis's authenticity.

- Data Consistency Assurance:

Consistency was a non-negotiable criterion. Efforts were concentrated on ensuring categorical values across the dataset were uniformly formatted, fostering a coherent analytical environment.

- Strategic Feature Engineering:

A significant enhancement was the introduction of the 'growth duration' attribute, a calculated feature that was extrapolated from the dataset's existing 'planting' and 'checked' dates. This new feature added a layer of depth to the analysis.

- Evaluating Class Distribution:

A balanced distribution among target classes is beneficial for modeling. The distribution of the growth categories was analyzed for each plant type, ensuring an even distribution. For instance, the following graph (Figure 2-12) showcases the growth category distribution for the 'Lemon Lime' plant, illustrating a largely balanced spread. Similar patterns were observed for other plant varieties.

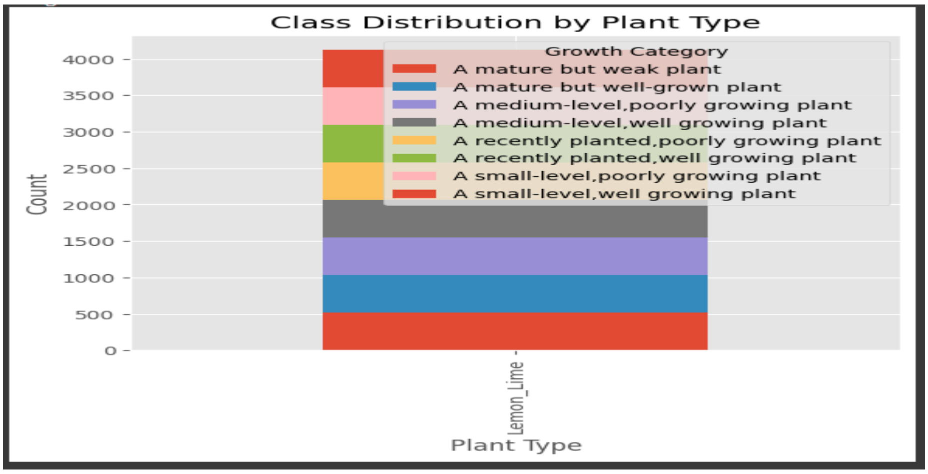


Figure 2-12 Growth category distribution by plant type

- Holistic Data Visualization and Dataset Segmentation:

Utilizing advanced visualization techniques, the dataset was portrayed in a comprehensive manner, facilitating a deeper understanding of the data's distribution. Following this, a strategic segmentation was executed, dividing the dataset into training and testing subsets in a calculated 80:20 ratio, ensuring ample data for both model training and subsequent evaluation.

- Categorical Data Processing:

To pave the way for advanced analytical techniques, categorical attributes underwent a transformation to numerical equivalents using the one-hot encoding methodology.

V. Model Training and Selection Process

A multiclass classification approach was deemed essential to address the complexities of the growth monitoring challenge. Various algorithms, each with its unique capabilities and advantages, were tested:

- Random Forest Classifier: With its reputation as a powerhouse in classification, this algorithm was deployed, culminating in a commendable accuracy of 0.96845.

- Decision Tree Classifier: Adopting an inherently intuitive hierarchical approach, this model was chosen for its straightforwardness, with the results reflecting an impressive accuracy of 0.97087.

- XGBoost Classifier: Tapping into the advanced gradient boosting domain, the XGBoost classifier, renowned for its unmatched performance, was put to the test. The results were indicative of its prowess, as it secured the highest accuracy of 0.97209.

- Bernoulli Naïve Bayes Classifier: Although binary-focused, its alignment with the dataset was found wanting, evident from its modest accuracy of 0.59345.

- LightGBM Classifier: A gradient boosting framework, LightGBM, renowned for its agility and effectiveness, was also tested. The resultant accuracy was a promising 0.96602.

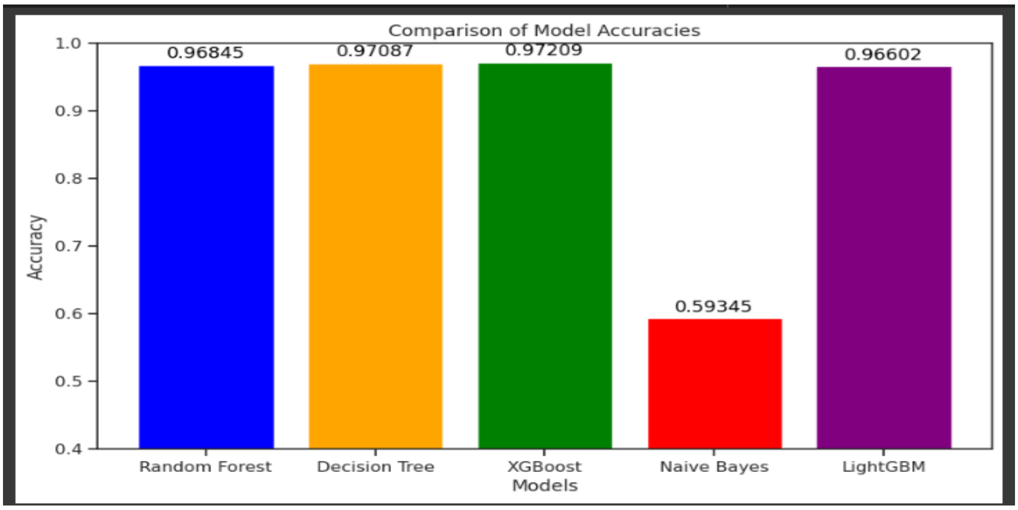


Figure 2-13 Comparison between the accuracy

Post the intensive training regimen, the models were put under the lens, subjected to a thorough evaluation (Figure 2-13). The emphasis was undeniably on accuracy, leading to the consensus that the XGBoost model, with its unparalleled precision, was best suited for this research challenge. This outcome underscored the essence of algorithmic compatibility and its paramount importance in deriving meaningful insights from data.

02.) Prediction System for Ornamental Plant Growth Duration

I. System Component Understanding:

An integral facet of the research was the design and deployment of a system to predict the time required for ornamental plants to attain their standard growth size. The urgency of this need arises in the commercial realm, where the time to maturity significantly influences order fulfillment and business dynamics. Sri Lanka's Floriculture Industry predominantly relies on manual expertise and experiential insights to navigate this challenge. This research aimed to digitalize and refine this prediction process. The primary objective was to ascertain the growth duration until a plant reaches the desired maturity for order placement, factoring in influential variables like growth medium, fertilizers employed, and the current growth phase.

II. Data Acquisition and Dataset Overview:

Incorporating empirical data, the research collaborated with Omega Green (PVT) Ltd. This association furnished access to comprehensive historical data, detailing specifications such as plant type, variety, growth medium, and fertilizer usage. Each plant's profile was built on critical attributes like plant variety, growth medium, fertilizer type, current pot type, and more. Though the original data was extensive, growth duration, a key attribute, was inferred from the existing dataset, enhancing the overall richness and relevance of the data for predictive purposes.

III. Class Attribute Determination:

The crux of the prediction system was to estimate the duration, or the number of days required for a plant to achieve its desired maturity. The derived attribute "days to order," representing this duration, became the cornerstone for predictions. Factoring in variables such as growth medium, fertilizer type, and shade type, the system aims to furnish precise estimates on the growth timeframes.

IV. Data Preprocessing and Cleaning:

Heralding the significance of pristine data, an exhaustive preprocessing routine was orchestrated. This commenced with understanding the dataset's structure and delineating correlations. Null values, both numerical and categorical, were identified and rectified, ensuring data integrity. Uniformity and standardization were imposed on categorical data. Pivotal feature engineering steps led to the introduction of ‘growth duration’ and transformation of planting and ordering dates into the more pertinent ‘days to order.’

- Evaluating Class Distribution

It's invariably advantageous for the modeling process to have an equitably distributed target class. The distribution of "days to order" was analyzed in this context. Observations revealed a slight rightward skew in the data, a not uncommon occurrence with real-world datasets.

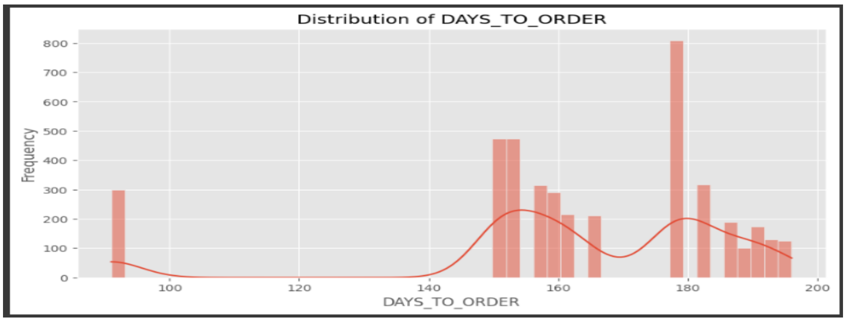


Figure 2-14 Before class distribution

Subsequently, a logarithmic transformation was applied to moderate the skewness. As shown in figure 2-15.

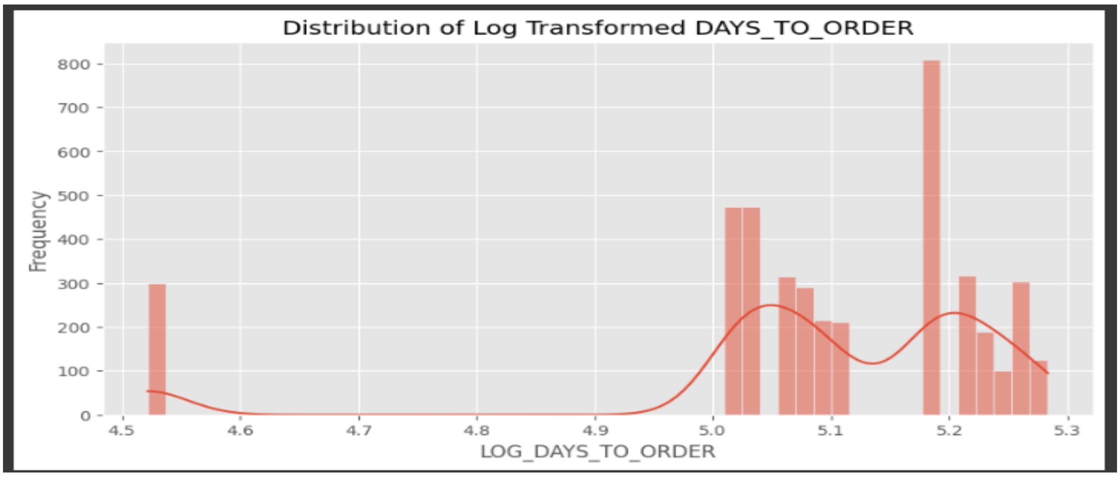


Figure 2-15 After logarithmic transformation

Then the strategic dataset split prepared the data for subsequent modeling.

V. Model Training and Selection:

Transitioning to the model training phase, the research ventured into regression models. Time-series algorithms, often the go-to for date forecasting, were bypassed in favor of regression models, prized for their ability to predict continuous outcomes like duration.

- Simple Linear Regression: Initial forays with simple linear regression hinted at potential misalignment, with the model reflecting a modest R2 score of 0.2505.

- Random Forest Regressor: Delving into ensemble methods, the Random Forest Regressor was harnessed. Known for its prowess in decoding non-linear relationships, it reflected commendable potential, achieving an R2 score of 0.9494.

- XGBoost Regressor: The gradient boosting marvel, XGBoost Regressor, was the final model of choice. Designed for efficiency and accuracy, it stood out, recording an R2 score of 0.9496 with an RMSE of 5.58, highlighting its aptness for this prediction challenge.

- LightGBM Regressor: The LightGBM Regressor was next in line for testing. A peculiarity was noted in its expectation for input data, specifically the need for inputs devoid of special characters. Post this modification, the model was optimized for regression and its performance gauged. The metrics underlined an R2 score of 0.9486, while the RMSE was recorded at 5.621.

- CatBoost Regressor: Concluding the model evaluations, the CatBoost Regressor was introduced into the fold. Its distinct ability to natively process categorical data, combined with its gradient-boosting capabilities, made it a promising candidate. On assessment, the model revealed an RMSE of 5.58, paired with an R2 score of 0.9495.

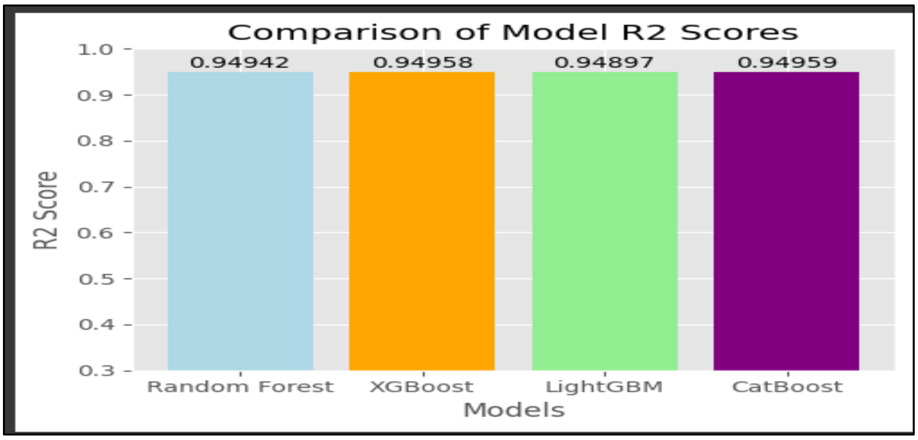


Figure 2-16 Comparison between R2 scores

Post rigorous evaluations, the XGBoost Regressor ascended as the optimal choice, shining in multiple facets. It championed an impressive R2 score of 0.94958, mirroring its adeptness in capturing dataset variance. Moreover, it brandished the lowest RMSE value of 5.58, pointing towards the precision of its predictions. Beyond accuracy, the XGBoost Regressor also marked high scores in computational efficiency, amalgamating rapid training times with uncompromised performance. This harmony of accuracy and efficiency, framed within the XGBoost Regressor, confirmed its suitability for navigating the intricacies of predicting ornamental plant growth durations.

2.2.4 Demand prediction system for ornamental plants

Data Collection:

The process of preparing for accurate demand forecasting in the floriculture industry involves several crucial steps to ensure that businesses can make informed decisions, efficiently allocate resources, and meet market demands with precision.

Identification of Relevant Data: The first step is to identify and select the necessary data components essential for demand forecasting. This includes historical sales data, order details, geographical distribution, seasonal variations, market trends, and customer behavior. The goal is to compile a comprehensive dataset that can facilitate precise demand predictions.

Determination of Time Frames and Sales Cycle Length: The second step involves defining the specific time frames for analysis, whether it's on a monthly, quarterly, or annual basis. Additionally, it entails determining the length of the sales cycle by considering factors like order lead times, production cycles, and order fulfillment duration.

Addressing Seasonality: Recognizing and accommodating seasonality is crucial in the floriculture industry as it significantly influences demand patterns. This step involves identifying peak and off-peak seasons, understanding their impact on product demand, and adapting forecasting models accordingly to ensure accurate predictions.
Data Collection and Organization: At the core of the process lies the systematic collection and organization of data. During this step, data elements identified earlier are collected, cleaned, validated, and structured in an organized manner. This well-

organized dataset serves as the fundamental input for developing robust demand forecasting models.

Implementing this structured approach to data collection and analysis is essential for the floriculture industry. It enables businesses to effectively predict demand, optimize resource allocation, and respond accurately to market dynamics. A well-prepared dataset, as shown in Figure 2-17, is the cornerstone of this process, ensuring that the industry can thrive in a dynamic and competitive landscape.

```
#Load the dataset
data = pd.read_csv('/content/drive/MyDrive/PP2/Data_set/DemandForcasting_final.csv')
```

+ Code + Text

```
data
```

	DateOfSale	Temperature	Humidity	Rainfall	Philo Sando	Aglonema	Lemon lime	Dendrobium	Ferns	Zamioculus	Bengamina	Dracaena	Country	PotSize	Season	Event
0	1/1/2012	27.5°C	81	227.57	8950	2410	3880	3620	2510	2280	2900	3690	France	5cm	Winter	IPM Dubai
1	1/2/2012	25.7°C	80	166.40	6930	4000	3110	2360	3270	3270	2970	2910	France	5cm	Winter	Japan Grand Prix International Orchid Festival
2	1/3/2012	26.3°C	78	147.26	2470	3380	8570	3530	2650	3290	2530	3290	France	8cm	Winter	IPM Dubai
3	1/4/2012	25.2°C	81	191.72	2420	2440	8880	2760	3570	3630	2670	2310	United Kingdom	5cm	Winter	IPM Essen
4	1/5/2012	27.8°C	77	236.65	3270	5990	3160	2750	2400	2570	3080	3640	Argentina	12cm	Winter	Japan Grand Prix International Orchid Festival

Figure 2-17 - Initial data set

Data Preprocessing and cleaning:

Missing Data Handling: In this initial step, we will rigorously assess the dataset for missing values. Missing data can adversely affect the accuracy of demand forecasting models. We will employ appropriate techniques, such as imputation or removal, depending on the nature and extent of the missing data. The choice of strategy will be data-driven and tailored to the specific dataset under examination.

Outlier Detection and Treatment: Identifying outliers is crucial as they can skew demand forecasting results. We will employ statistical methods and visualization techniques to detect outliers within the dataset. Once identified, we will decide on an appropriate strategy for handling outliers, such as transformation, imputation, or removal. The choice will be contingent upon the dataset's characteristics and the impact of outliers on the forecasting models. The following figures shows before and after of outlier removing. Figure 2 – 18 and 2 – 19 showing the outliers checking before and after.

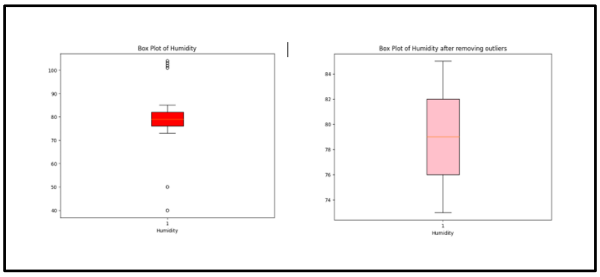


Figure 2-19 - Outliers checking before and after humidity

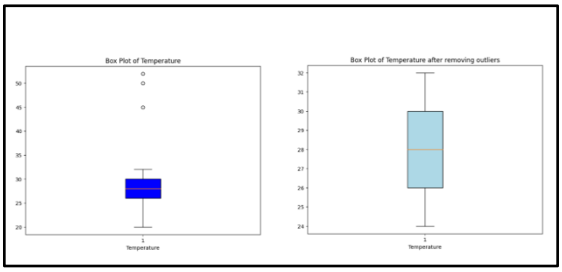


Figure 2-18 - Outliers checking before and after for temperature

Data Quality Assessment: Ensuring data quality is paramount for robust forecasting. We will scrutinize the dataset for potential data quality issues, including duplicate entries, inconsistent formatting, and data entry errors. Any identified issues will be meticulously addressed to enhance data integrity.

Data Transformation: To facilitate effective analysis, we will transform the data into a suitable format. This transformation may involve converting categorical variables into numerical ones, especially if the chosen forecasting model relies on numerical inputs. The specific transformation techniques will be determined based on the forecasting model to be applied and the dataset's requirements.

The data preprocessing steps outlined here are integral to the success of demand forecasting within the floriculture industry. By systematically addressing missing data, outliers, data quality issues, and data transformation, we will ensure that the dataset is ready for rigorous analysis. These preparatory steps are vital for achieving accurate and actionable demand predictions, contributing to the overall success of the research and its practical applications within the industry.

The transformed data will be done in the model building as per Figure 2-9. And additionally, one hot encoding was used as a preprocessing technique for categorical variables while numerical data handled by scaling using minmax scaler.

Feature selection and feature engineering:

Identification of Relevant Features: In the first step, we will identify a comprehensive set of features that may have an impact on demand within the floriculture industry. These features may include:

Weather Data: Information about temperature, humidity, precipitation, and other meteorological factors that can influence floral growth and customer demand.

Holidays and Events: Data on holidays, festivals, and special events, as these occasions often drive fluctuations in demand for floral products.

Market Trends: Data on broader market trends, such as economic indicators, consumer sentiment, and floral design preferences, which can affect demand patterns.

Exploratory Data Analysis (EDA) and Correlation Analysis: EDA techniques will be employed to gain a deeper understanding of the dataset and its features. We will visualize the data to identify patterns and trends. Additionally, we will perform correlation analysis to assess the relationships between different features and the target variable (demand). This analysis will help us pinpoint the most influential features.
Creation of New Features: To capture complex patterns and trends in the data, we will create new features. These may include:

Lagged Variables: By introducing lagged versions of the target variable or other relevant features, we can account for time dependencies and past trends in demand.

Rolling Averages: Calculating rolling averages or moving averages can help in smoothing out noise in the data and identifying long-term trends.

Seasonality Indicators: Creating binary or categorical indicators for different seasons or specific time periods can help model seasonal variations in demand.

Feature engineering is a crucial aspect of demand forecasting within the floriculture industry. By identifying relevant features, conducting exploratory data analysis, and creating new features, we aim to enrich the dataset and enhance the predictive power of our models. These steps are essential for capturing and modeling the intricate demand patterns and drivers in the floriculture sector, ultimately leading to more accurate and actionable forecasts. Handling Imbalance with Stratified Sampling: Finding any class imbalance in the target variable, such as an unequal distribution of demand across various products or regions. Use stratified sampling to ensure that each class is represented in training and testing datasets in proportion to their overall frequency. The figure 2 – 21 and 2 – 20 showing the frequencies and distributions accordingly.

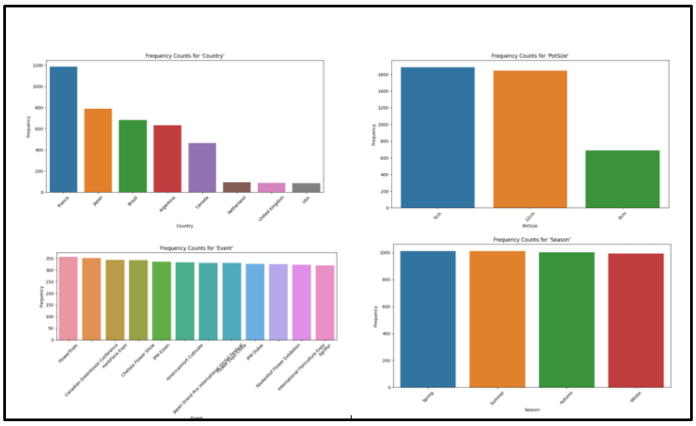


Figure 2-21 - Frequencies of categorical variables

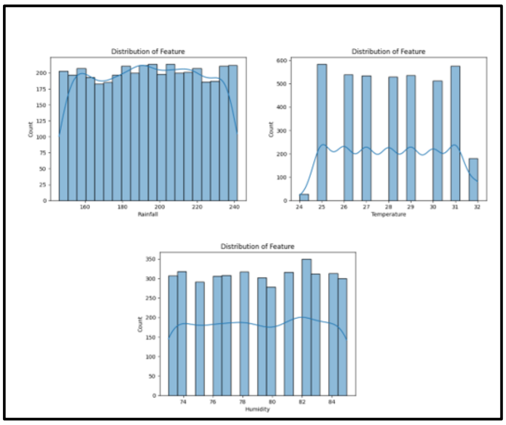


Figure 2-20 - Distributions of numerical variables

Splitting the data set:

In the process of preparing data for demand forecasting in the floriculture industry, three key steps are employed. Firstly, the dataset is separated into training (70% of the data), testing (30% of the data), and a separate validation set. Secondly, random sampling and stratification techniques are applied to ensure the subsets are representative and account for any imbalances or patterns. Additionally, feature scaling techniques like MinMaxScaler are applied to normalize the data, ensuring that all features contribute effectively to the models without any dominating the analysis.

Model Building:

The process of selecting the optimal demand prediction model in the floriculture industry involves four key steps. Firstly, a range of models, including Linear Regression, Random Forest, Gradient Boosting, MLP Regressor, and SARIMA, is considered. Second, these models are trained using the training dataset to learn from historical demand data and features. Third, their performance is rigorously evaluated on the validation dataset using metrics like R-squared, Mean Squared Error, and Mean Absolute Error. Finally, the model that performs best based on these metrics is chosen as the most suitable for accurate demand forecasting. This meticulous model selection and evaluation process ensures that the floriculture industry can rely on a robust predictive model to make informed decisions and optimize operations effectively.

Model Accuracy and Selecting the Best Model:

In the final stages of demand prediction for the floriculture industry, the chosen model's performance is rigorously evaluated using a testing dataset, with metrics like R-squared, Mean Squared Error (MSE), and Mean Absolute Error (MAE), and iterative improvements are made if necessary, resulting in the selection of Linear Regression as the final model due to its transparency and interpretability. The best model selected as from the Figure 2 – 22.

```
Best models for each target column:
sqrt_Philo Sando    Gradient Boosting
sqrt_Aglonema       Linear Regression
sqrt_Lemon lime     Linear Regression
sqrt_Dendrobium     Linear Regression
sqrt_Ferns          Linear Regression
sqrt_Zamioculus     Linear Regression
sqrt_Bengamina         Neural Network
sqrt_Dracaena       Linear Regression
dtype: object
```

```
1 # the best approach is Linear regression
```

Figure 2-22 - Models for each plant type

2.3 Integration Process

The integration process seamlessly connects all backend components through APIs, ensuring smooth communication. These APIs are purposefully designed to receive HTTP POST requests, efficiently process the data within them, and provide responses in JSON format. Flask serves as the web framework, Python as the programming language, and PyCharm as the chosen IDE for API development.

On the front-end side, React Native is utilized within Visual Studio Code to create a cohesive user experience. The four components are intricately linked through routes, allowing users to navigate effortlessly. User inputs are collected from the mobile app's frontend, then relayed to the backend API for preprocessing. The data is processed using the model, and the resulting output is seamlessly transmitted back to the mobile app's frontend, where it is presented to the user in a user-friendly format. This intricate system ensures a streamlined and efficient flow of information between the user and the application's backend components.

2.4 Visual Flow of Frontend

Upon its launch, the application presents users with a brief splash screen before transitioning to the home screen, which acts as the central navigation hub. Throughout the application, a consistent design theme is applied to all navigation headers, featuring a distinctive font size, bold typography, and a unique color scheme to maintain visual coherence. This central hub provides access to a range of functionalities, granting users the ability to select from four options as in Figure 2-12.

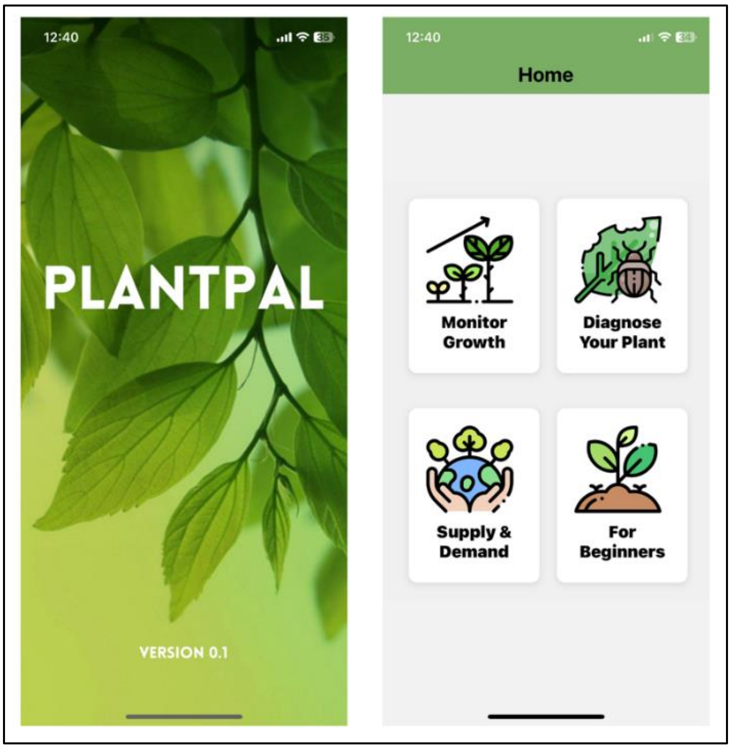


Figure 2-23 User interfaces of Plant pal

There are two primary types of user inputs in the application:

1) Form Inputs:
   Users can input relevant details into a form and then tap the button below the form to receive predictions. Figure 2-13 provides an example of the form interface where users can input their information.

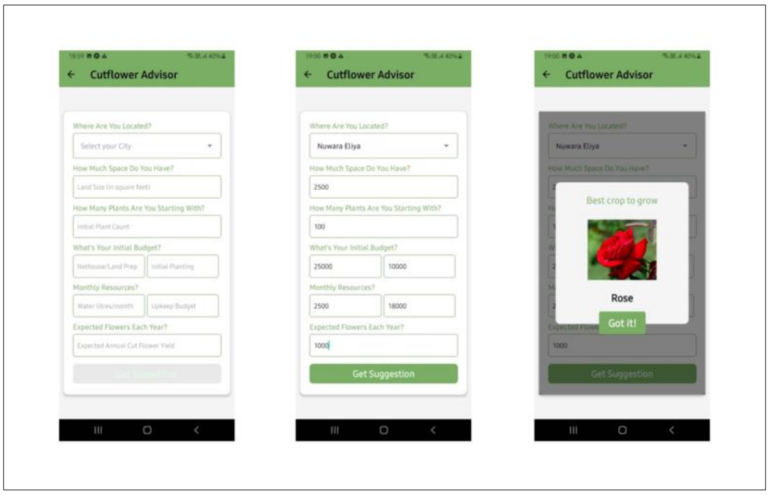


Figure 2-25 Form inputs

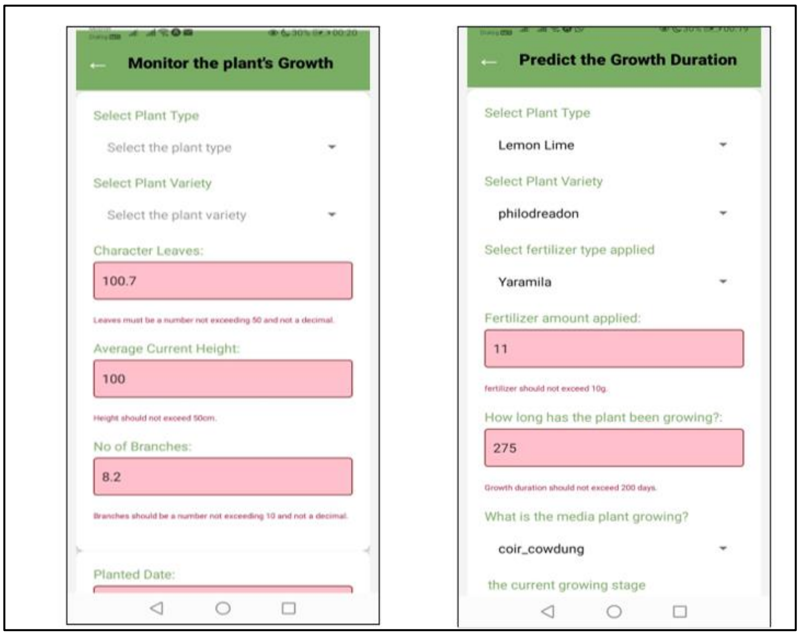


Figure 2-24 Form input validations

2) Image Inputs:

Users have the option to capture real-time images using their smartphone's camera or upload images from their gallery for predictions. Figure 2-14 illustrates an example of this image input feature.

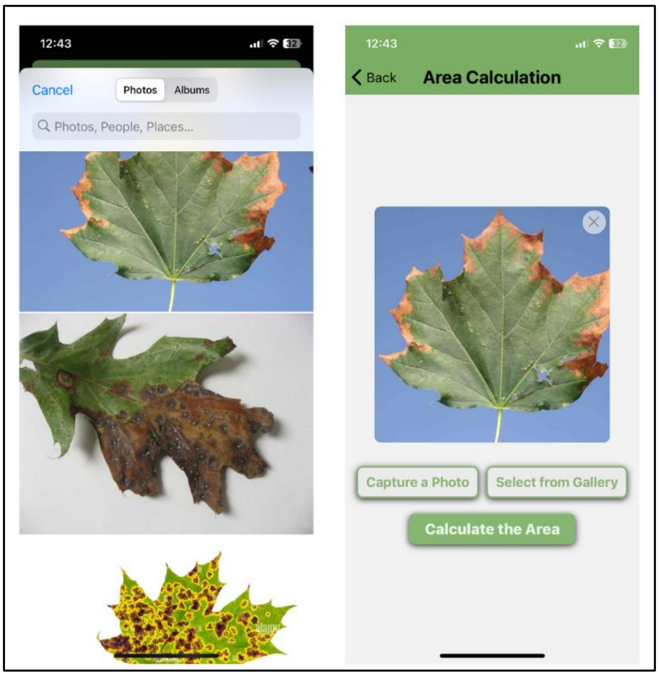


Figure 2-26 Image inputs

It's important to note that all user inputs undergo validation to ensure the delivery of reliable services. Figure 2-15, for instance, showcases an alert that appears when a user inputs an invalid image.

Figure 2-27 Validation alert

3 RESULTS AND DISCUSSION

3.2 Testing and Results

3.2.1 Leaf affliction analysis and shipment recommendation system

Subsequently, the application evaluates if the detected object in the image is likely a leaf based on certain criteria. If deemed a leaf, the application calculates the affliction percentage by segregating the healthy and afflicted areas. The results, along with images showcasing the leaf on a gray background, healthy areas, afflicted areas, and the mask, are stored in a designated 'uploads' directory. The paths to these processed images are then returned in the response payload, allowing the user to view or download them. The following Figure 3-1 shows how the images saved after the process.

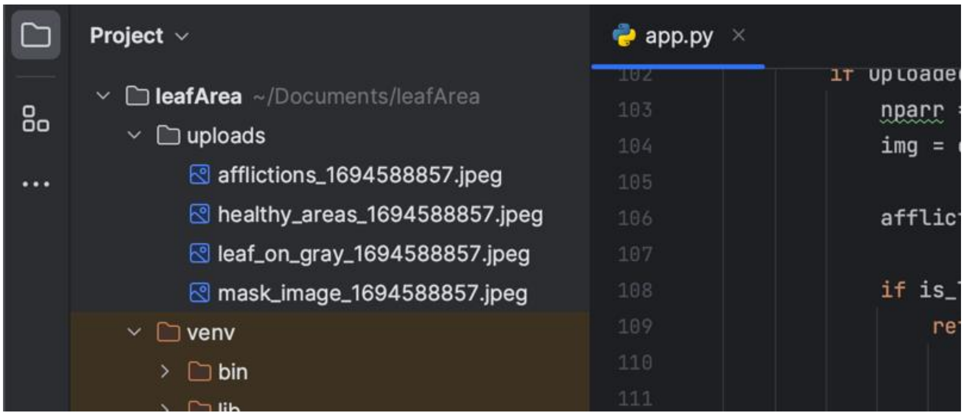


Figure 3-1 Saved images in the uploads folder

Once the image is processed, the backend generates various renditions: an image with the background removed, the healthy portion of the leaf, the afflicted regions on the leaf, and a mask image. The Figure 3-2 below showcase the backends' processing of both captured and gallery-selected images.

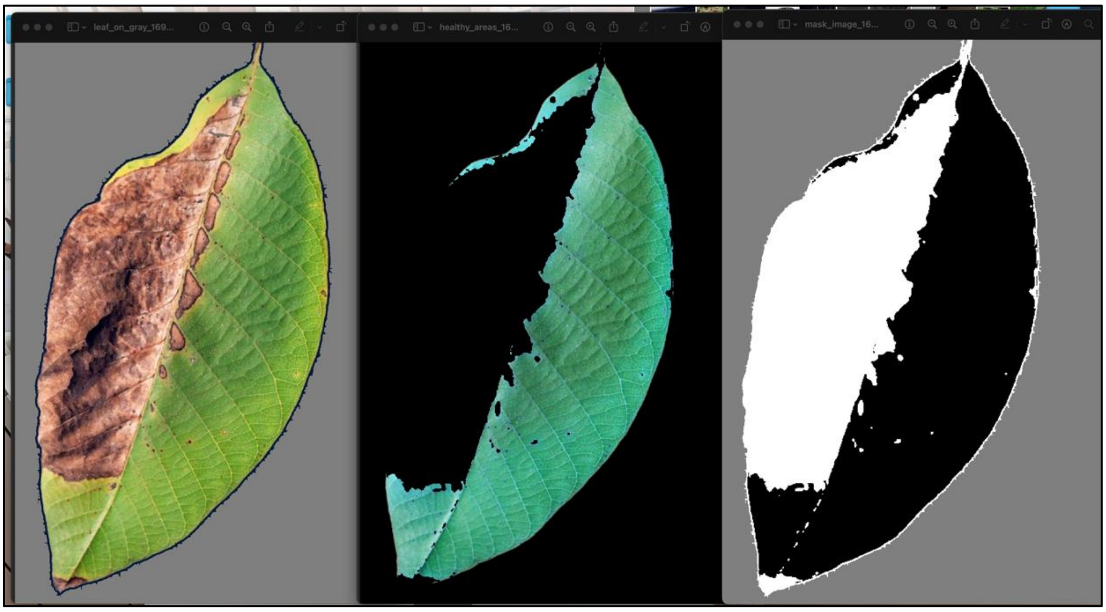


Figure 3-2 Backend generated masks

Figure 3-3 illustrates how the final result is visible to user in mobile application.

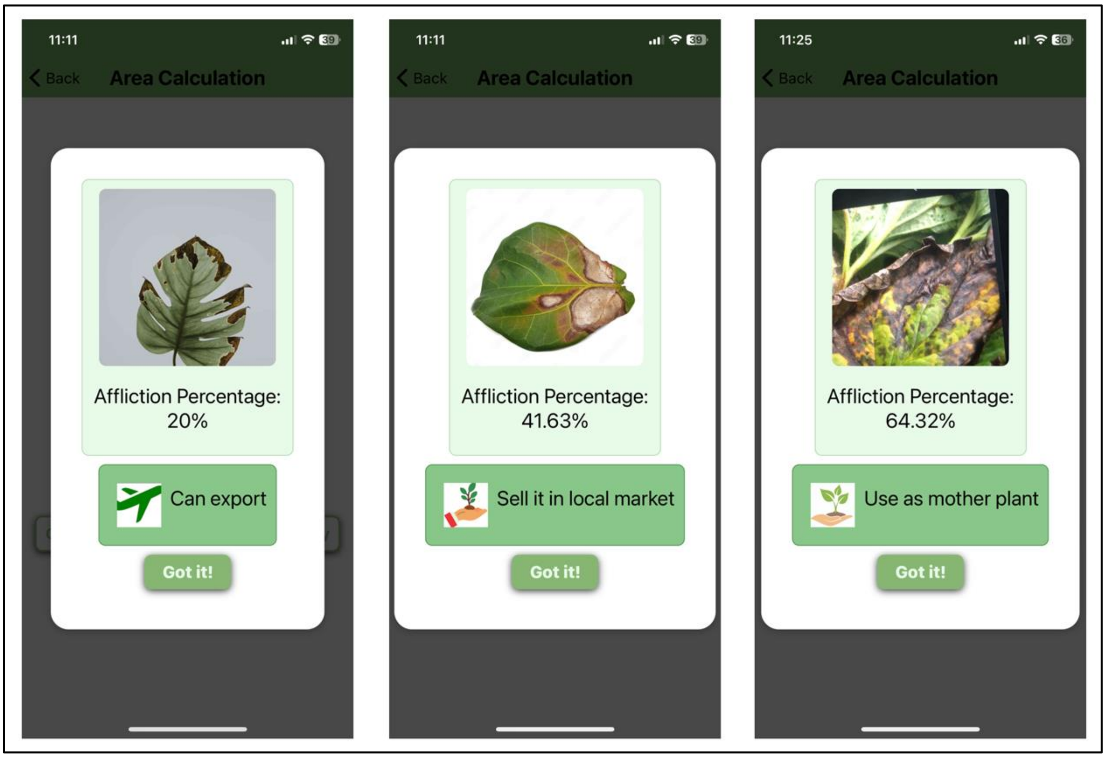


Figure 3-3 Results after calculation

3.2.2 Cut flower selection and philodendron variety identification system

The cut flower selection model scored a high accuracy score of 99.87%. 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 confusion matrix in Figure 3-4 demonstrates how well the model predicts unseen data.

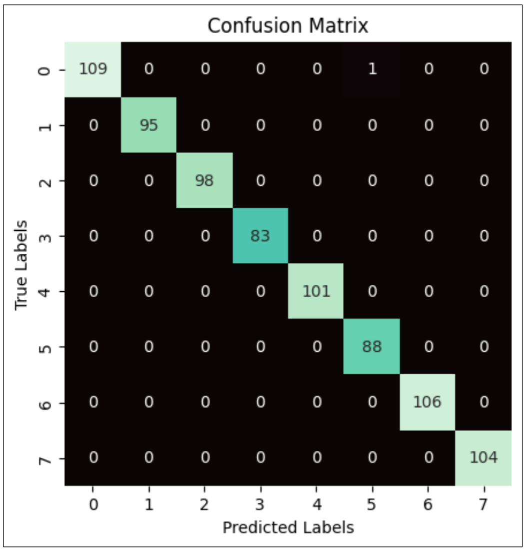


Figure 3-4 Cut flower selection confusion matrix

The below Figure 3-5 displays how the prediction is shown to the user.

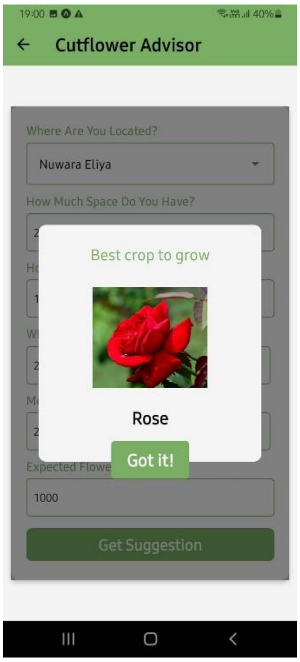


Figure 3-5 Cut flower selection result

The Philodendron classification model and the binary classification model exhibit good accuracy and minimal loss. The Philodendron classification model showed 82.89% accuracy and binary classification model showed 100% accuracy on unseen data. A sample of the test images is displayed along with the actual classes, predicted classes, and associated probabilities. Figure 3-6 shows the results for the sample testing set of the Philodendron classification model. Figure 3-7 shows the results for the sample testing set of the binary classification model.

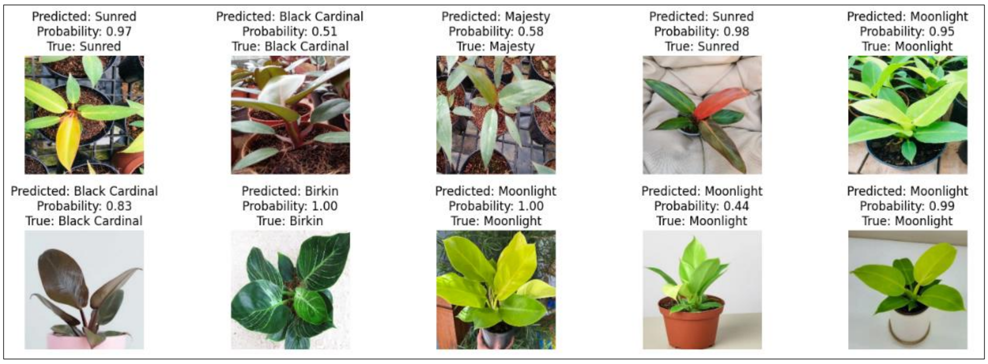


Figure 3-6 Philodendron classification model sample testing

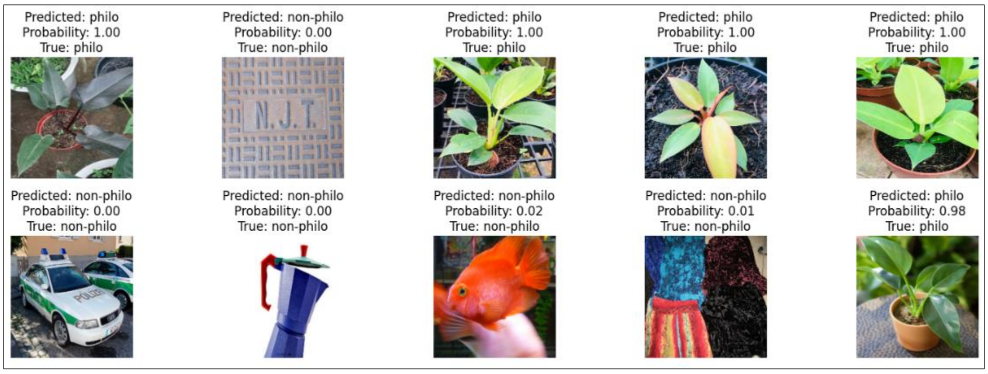


Figure 3-7 Binary classification model sample testing

Figure 3-8 illustrates how the identified philodendron variety is shown to the user.

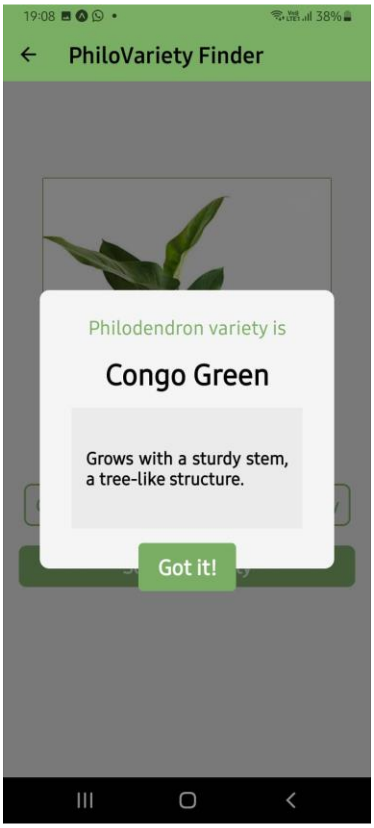


Figure 3-8 Philodendron classification result

3.2.3 Growth monitoring and prediction system

1) Growth Monitoring and Classification System for Ornamental Plants:

- Model Selection and Evaluation:

Among various models experimented with, the XGBoost classifier distinctly emerged as the frontrunner in the growth monitoring and classification segment. The XGBoost classifier, an embodiment of the advanced gradient boosting algorithm, is reputed for its rapid execution and unparalleled performance. This classifier was meticulously tailored for the project, considering its adeptness at managing multiclass classification challenges. The set objective of 'multi:softprob' enhanced its capability in this context.

It's imperative to underscore that the XGBoost algorithm demanded the dataset to be strictly numerical. Addressing this, the class variable underwent label encoding, ensuring seamless compatibility with the model's specifications.

A landmark achievement for the XGBoost classifier was its outstanding accuracy rate, peaking at 0.97209. This statistic not only validates the model's efficacy but also establishes it as an industry-standard tool for similar challenges.

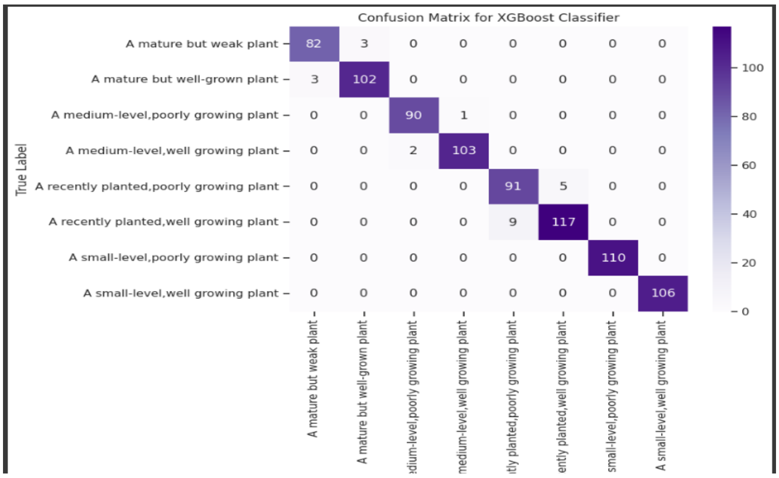


Figure 3-9 Confusion Matrix of XGBoost Classifier

- Confusion Matrix Analysis:

Post-evaluation, a confusion matrix was generated, offering insights into the model's performance nuances. The matrix clearly delineated the true positives, true negatives, false positives, and false negatives, presenting a comprehensive view of the model's prediction capabilities. The class attributes were inversed transformed to accentuate clarity and user comprehension in the displayed results. As illustrated in the subsequent figures, the confusion matrix substantiates the XGBoost classifier's prowess in predicting the current growth category of ornamental plants.

- User Interface Visualization:

Following the backend successes, it was also essential to evaluate the model's performance from an end-user's standpoint. Upon accurate input validation and submission by the user, the system promptly returns the plant's current growth category. The interface, as showcased in the accompanying figures, is designed for simplicity and user-friendliness, ensuring even individuals without a technical background can effortlessly navigate and derive value.

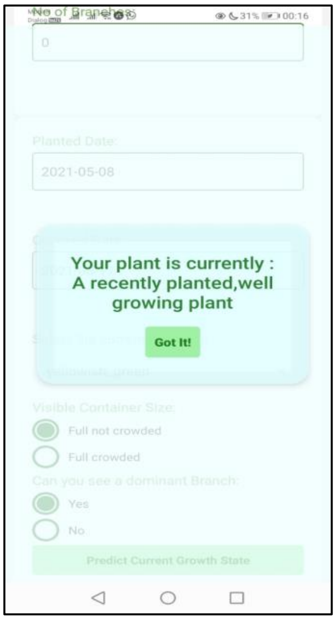


Figure 3-10 Prediction output of Growth Monitoring

In conclusion, the growth monitoring and classification system, powered by the XGBoost classifier, stands as a testament to the synergy of domain knowledge, advanced algorithms, and user-centric design. It not only streamlines the ornamental plant growth monitoring process but also propels the industry towards a more data-driven, efficient future.

02.) Prediction System for Ornamental Plant Growth Duration

- Model Selection and Performance:

In the quest to design an optimal prediction system for the growth duration of ornamental plants, the XGBoost Regressor was identified as the most potent tool. XGBoost, an avant-garde gradient boosting framework, is distinguished by its remarkable efficiency and precision in model predictions. To tailor the model specifically for this project, a series of parameters were judiciously set. The 'reg:squarederror' objective was particularly crucial as it ensured the model was attuned to regression tasks. The other parameters, namely 'n_estimators' and 'seed', were fine-tuned to bolster the model's regression prowess and cement reproducibility, respectively.

A notable accomplishment of the XGBoost Regressor was its achievement of an R2 score of 0.9496. This score not only indicates a near-perfect fit of the model to the data but also means that approximately 95% of the variability in the dataset can be explained by the model. Furthermore, the RMSE (Root Mean Square Error) of 5.58 serves as a testament to the model's accuracy, indicating minimal deviations from the true values in its predictions.

- Precision-Error Plot Analysis:

To further delve into the nuances of the model's performance, a precision-error plot was generated as shown in figure 3-11. This plot offers a visual representation of the model's prediction accuracy, juxtaposing predicted values against the actuals. The close alignment between the two, as portrayed in the subsequent figures, emphatically underscores the model's proficiency. The alignment is especially evident around the line of identity, which symbolizes perfect prediction. Deviations from this line were minimal, showcasing the model's commendable precision and reliability in predicting growth durations.

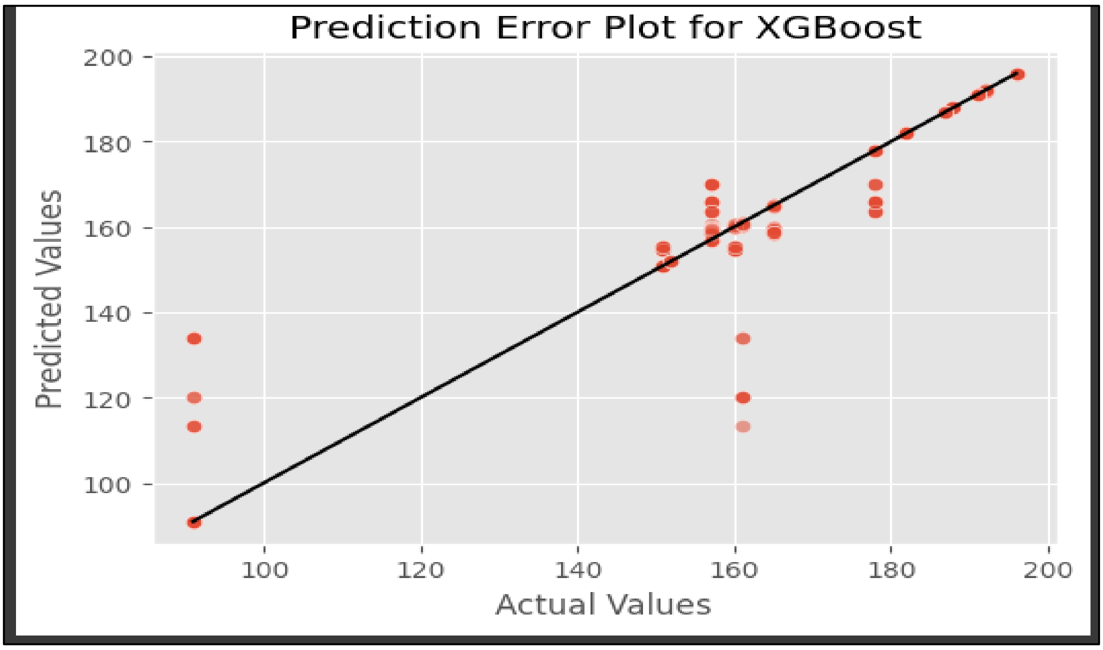


Figure 3-11 Prediction vs Actuals plot

- User Interface Visualization:

The culmination of the backend modeling process was realized through an intuitive user interface. This interface not only offers insights into the predicted growth duration but is also designed to be straightforward, ensuring ease of use. Upon input validation and submission, the system promptly returns an estimate of the number of days required for the plant to reach the desired growth maturity. This user-centric design, as showcased in the accompanying figures, enables stakeholders, ranging from agronomists to business managers, to gain actionable insights effortlessly.

To encapsulate, the XGBoost-powered prediction system for ornamental plant growth duration is a monumental stride towards harnessing data-driven insights in the floriculture industry. By accurately forecasting growth durations, stakeholders can optimize their operations, ensuring timely order fulfillments and maximizing profitability. This system epitomizes the perfect blend of domain expertise, technological innovation, and user-centric design, driving the industry into a future brimming with possibilities.

3.2.4 Demand prediction system for ornamental plants

The results derived from the meticulous testing of our Backend API in Postman play a pivotal role in the development and refinement of our demand forecasting system, a core component of the 'Plant Pal' application. This testing phase is strategically designed to focus on the accuracy and efficacy of our demand prediction capabilities, as this aspect is paramount for providing our users with precise and valuable insights into their plant care needs.

As we delve into the testing process, we thoroughly examine how well the API handles various user inputs, including plant types, temperature, humidity, rainfall, and pot sizes – all critical factors that influence demand forecasts. These tests rigorously evaluate the system's ability to process this information and generate forecasts that align with real-world conditions.

The results obtained from this demand forecasting testing serve as a crucial checkpoint in the evolutionary journey of our 'Plant Pal' application. They offer a tangible assurance of the API's functionality, confirming that it can accurately interpret user input, analyze data, and respond with meaningful forecasts. Furthermore, these results are invaluable in terms of quality assurance, as they allow us to identify and promptly address any potential issues or discrepancies in the demand forecasting process.

Figure 3 – 12 shows the inputs for the Rainfall, Temperature, Humidity, and the pot size which will give the prediction as the result of demand accordingly.

Ultimately, the demand forecasting testing conducted in Postman reaffirms our unwavering commitment to delivering a robust and dependable solution to our users. By ensuring that the API performs seamlessly, providing accurate demand forecasts, we enhance the value of 'Plant Pal' for both plant enthusiasts and business owners. It's through this rigorous testing that we continue to refine our system, assuring its reliability and effectiveness in assisting users with their plant care and business decisions.

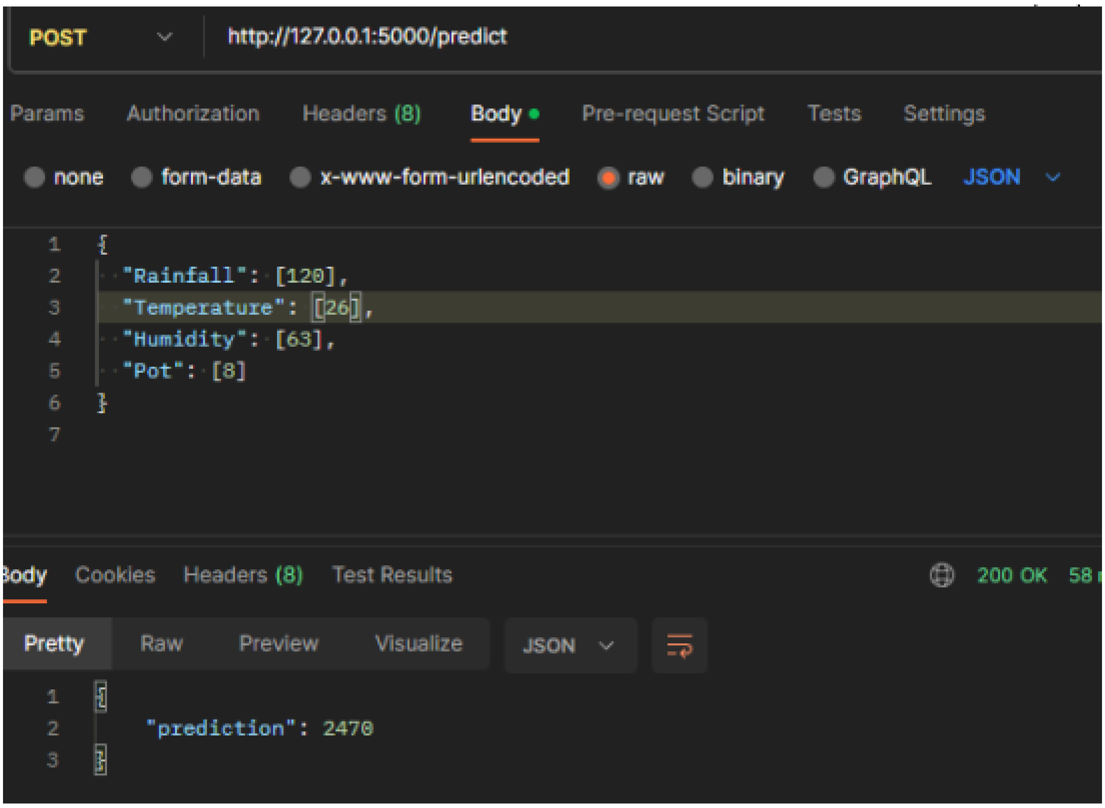


Figure 3-12 - Postman test results

3.3 Research Findings

Upon conducting our study, we discovered that the proposed mobile application functions with significant accuracy and efficiency in the floriculture context of Sri Lanka. Notably, the plant growth prediction tool demonstrated remarkable precision in forecasting the developmental trajectory of ornamental plants, while the leaf affliction analysis system accurately identified and quantified afflictions, streamlining the decision-making process for shipments. Furthermore, the plant recommendation and variety identification system proved invaluable for industry beginners, consistently providing informed guidance. The supply-demand forecasting tool also exhibited consistent reliability, aiding producers in their market assessments. Overall, our findings affirm the effectiveness and accuracy of the mobile application's functionalities, highlighting its potential to revolutionize the Sri Lankan floriculture sector.

3.4 Discussion

The findings of our research provide compelling evidence that technology, particularly machine learning and computer vision, can profoundly transform traditional sectors like floriculture in Sri Lanka. As our results indicate, the integration of advanced tools into an accessible mobile application bridges significant gaps that have long hindered efficiency and growth in the industry.

One of the paramount challenges in the floriculture sector has been the accurate prediction of plant growth. Traditionally, relying on manual methods and past experiences often resulted in imprecise predictions, affecting the overall yield and quality of the plants. Our research showed that with the help of the growth prediction tool embedded in the mobile application, users can forecast the growth trajectory with a precision that surpasses traditional methods. This is a game-changer for producers, allowing for better resource allocation and optimization of cultivation techniques.

Equally impactful was the leaf affliction analysis system. As highlighted in the findings, this system could potentially reduce financial losses by aiding growers in ensuring their shipments meet export standards. In a sector where the difference between a successful shipment and a rejected one can be minute visual differences in plant health, having a tool that can provide accurate affliction readings is invaluable.

For beginners, the world of floriculture can be overwhelming. The intricacies of choosing the right plant variety, understanding its needs, and predicting its market demand can deter many from entering the industry. Our application’s plant recommendation system, backed by data-driven insights, ensures that newcomers are not left to grapple with these complexities unsupported. By offering guidance on plant varieties suitable for specific conditions and market demands, the app has the potential to reduce the steep learning curve for newcomers and, in turn, increase the number of successful floriculture ventures.

The development of a comprehensive demand forecasting system for the floriculture industry involves a systematic approach, starting with data collection and preprocessing to address data quality issues. Machine learning models, including Linear Regression is selected to optimizing operations and enabling data-driven decision-making in the floriculture sector.

However, while our research shows promising results, it's crucial to understand that the application's effectiveness will also depend on user engagement and adaptability. Future studies should focus on the app's long-term impacts, user feedback, and potential areas for improvement.

4 SUMMARY OF EACH STUDENT'S CONTRIBUTION

Table 4-1 Individual contribution

Student ID	Student Name	Contribution
IT19994406	Basnayake N.S.N.	• Calculate the leaf afflictions in percentage. • Build the recommendation system for export or not. • Validate the invalid images. • API development leaf affliction part. • UI designing for the leaf affliction part. • Build the mobile application for both IOS and android to calculate the leaf afflictions. • Finding the correct HSV ranges for the healthy leaf. • Participating in UI designing • Integrate the mobile app
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 • Integrate the mobile app
IT20005726	Liyanage S.R.	• Identified plant growth using features like leaf color, height, and number of leaves. • Developed a system to monitor the growth of Sri Lanka's popular ornamental plants. • Designed a system to forecast how long plants need to reach their desired growth based on current conditions and environmental factors. • Created a cross-platform mobile app using React Native for growth tracking and predictions. • Developed a user-friendly interface that accurately validates user inputs. • Tailored the research to address practical challenges in the floriculture industry. • Contributed to the design and integration of the user interface.
IT20017088	Prabhashi P.A.N.	• Obtain the dataset using the manually inputted data from the client. • Using the manually developed data set select the best variables for the model building.

		• Used different kinds of preprocessing techniques for higher accuracy. • Using the best fitted model build the backend. • Using the API validate the test results taken from the backend. • Cross platform mobile application developing for the individual component. • Participating the group developments as required.

5 CONCLUSION

The floriculture sector in Sri Lanka presents a remarkable economic opportunity, but it is hindered by a myriad of challenges stemming from the heavy reliance on outdated technologies and manual processes. For newcomers, the intricate web of plant variety selection, growth monitoring, adherence to strict export standards, and the dire need for accurate plant identification complicates their entry into the industry. These challenges are exacerbated by the current lack of an efficient and comprehensive system to guide and assist them.

This research has aimed to address these significant gaps in the sector by leveraging advanced technologies such as machine learning, deep learning, and computer vision. The envisioned mobile application is set to revolutionize the way individuals, especially beginners, approach floriculture. By offering functionalities like plant growth prediction, leaf affliction analysis, plant recommendation, and supply-demand forecasting, the application promises not just to simplify tasks but also to ensure that decisions are backed by data-driven insights.

The requirements laid out for this system underscore the need for accuracy, user-friendliness, efficiency, and adaptability. The combination of Google Colab, TensorFlow, Keras, Scikit-Learn, OpenCV, Flask, and React Native not only ensures that the application is built on a robust technological foundation, but it also guarantees that users have a seamless experience, be it capturing photos, uploading images, or receiving recommendations.

In sum, the importance of this research and the resulting application cannot be overstated. By providing a holistic solution that caters to both the functional and non-functional requirements of the floriculture sector, this project holds the promise of significantly transforming the industry. It can reduce the steep learning curve for newcomers, optimize processes for seasoned professionals, and overall elevate the quality, efficiency, and profitability of floriculture in Sri Lanka. As the world continues its rapid technological evolution, it is projects like these that will ensure traditional sectors such as floriculture are not left behind but are instead propelled forward into a brighter, data-driven future.

REFERENCES

[1] M. o. I. Promotion, "FLORICULTURE PRODUCTS," Sri Lanka Export Development Board, [Online]. Available: https://www.srilankabusiness.com/floriculture/overview.html. [Accessed 30 01 2023].

[2] 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.

[3] 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.

[4] M. U. A. B. A. R. H. T. E. F. e. a. Matteo G. Ziliani, "Early season prediction of within-field crop yield variability by assimilating CubeSat data into a crop model," *Agricultural and Forest Meteorology*, vol. 313, p. 108736, 2022.

[5] A. S. G. G. &. B. D. Paturkar, "Plant trait measurement in 3D for growth monitoring," *Plant Methods 18*, vol. 18, no. 1, 2022.

[6] B. I. a. R. E. S. R. Ubudi, "Automation system for controlling and monitoring ornamental plants using fuzzy logic method," in *International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC)*, Yogyakarta, Indonesia, 2017.

[7] P. S. D. a. B. C. B. A. Saha, "Smart Green House for Controlling & Monitoring Temperature, Soil & Humidity Using IOT," in *2nd International Conference on Artificial Intelligence and Signal Processing (AISP)*, Vijayawada, India, 2022.

[8] O. E. C. Y. Kirci P, "A Novel Approach for Monitoring of Smart Greenhouse and Flowerpot Parameters and Detection of Plant Growth with Sensors," *Agriculture*, vol. 12, no. 10, p. 1705, 2022.

[9] D. K. S. D. R. P. Sahana B, "Smart Green House Monitoring based on IOT," *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT),* vol. 8, no. 14, 2020.

[10] C. &. K. S. D. &. K. K. &. S. K. &. V. S. &. J. S. Visvesvaran, "Smart Greenhouse Monitoring System using Wireless Sensor Networks.," in *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*, 2021.

[11] Y.-S. a. L. T.-H. a. Y. K. S. Tong, "Deep Learning for Image-Based Plant Growth Monitoring: A Review," *International Journal of Engineering and Technology Innovation,* vol. 12, pp. 225-246, 2022.

[12] M. X. S. Y. Y. J. D. S. P. Yao Meng, "Flexible and high quality plant growth prediction with limited data," *Frontiers in Plant Science,* vol. 13, 2022.

[13] A. &. P. S. Suthakaran, "Detection of the affected area and classification of pests using convolutional neural networks from the leaf images," *International Journal of Computer Science Engineering and Information Technology (IJCSE),* vol. 9, 2020.

[14] V. K. S. a. R. P. A. Bhattacharjee, "Localization and Estimation of Plant Disease Intensity for Aerial Image," in *International Conference on Intelligent Computing and Remote Sensing (ICICRS)*, Bhubaneswar, India, 2019.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] D. H. L. A. M. C. K. A. B. L. a. F. W. M. K. S. S. W. U. S. S. Samaratunge Arachchillage, "Smart Intelligent Floriculture Assistant Agent (SIFAA)," in *3rd International Conference on Advancements in Computing (ICAC)*, Colombo, Sri Lanka, 2021.

[21] K. K. T. S. D. D. A. K. a. N. P. S. Kulasekara, "Machine Learning Based Solution for Improving the Efficiency of Sugar Production in Sri Lanka," in *13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, 2022.

[22] H. R. A. S. G. H. J. W. a. D. N. R. Gamage, "Smart Agriculture Prediction System for Vegetables Grown in Sri Lanka," in *IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, 2021.

[23] B. V. B. P. a. M. Dakshayini, "Demand-prediction model for forecasting AGRI-needs of the society," in *International Conference on Inventive Computing and Informatics (ICICI)*, Coimbatore, India, 2017.

[24] F. A. D. M. N. W. W. S. A. R. a. R. R. Ismayanti, "Identification of Pests on Black Orchid Plants Using Naïve Bayes Method Based on Leaf Image Texture," in *2022 International Conference of Science and Information Technology in Smart Administration (ICSINTESA)*,, Denpasar, Bali, Indonesia, 2022.

[25] M. A. H. I. H. S. S. A. I. J. L. Hasan Ahmed, "Detection and classification of plant diseases in leaves through machine learning," *Indonesian Journal of*

Electrical Engineering and Computer Science (IJEECS), vol. 28, no. 3, pp. 1676-1683, 2022.

[26] A. J. a. S. N. a. V. U. a. D. Viswanath, "Detection and classification of nutrient deficiencies in plants using machine learning," *Journal of Physics: Conference Series,* vol. 1850, 2021.

[27] M. M. R. M. J. W. a. A. S. Y. Sari, "Detection of Corn Leaves Nutrient Deficiency Using Support Vector Machine (SVM)," in *2021 4th International Conference of Computer and Informatics Engineering (IC2IE)*, Depok, Indonesia, 2021.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] M. U. A. B. A. R. H. T. E. F. Y. L. J. S. I. H. M. F. M. Matteo G. Ziliani, "Early season prediction of within-field crop yield variability by assimilating

CubeSat data into a crop model," 11 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168192321004226. [Accessed 28 03 2023].

[34] B. I. a. R. E. S. R. Ubudi, "Automation system for controlling and monitoring ornamental plants using fuzzy logic method," in *International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC),* Yogyakarta, Indonesia, 2017.

[35] G. B. e. al., "Information technology controlled greenhouse: A system architecture," in *2018 IoT Vertical and Topical Summit on Agriculture - Tuscany (IOT Tuscany),* Tuscany, Italy,, 2018.

[36] Y.-S. &. L. T.-H. &. Y. K. S. Tong, "Deep Learning for Image-Based Plant Growth Monitoring: A Review," *International Journal of Engineering and Technology Innovation,* vol. 12, no. 3, pp. 225-246, 2022.

[37] G.-Y. &. L. R.-R. &. C. Z. &. T. K. &. S. M.-F. &. Y. G.-H. &. L. Q.-Q. Zhang, "Manipulation of the Rice L-Galactose Pathway: Evaluation of the Effects of Transgene Overexpression on Ascorbate Accumulation and Abiotic Stress Tolerance," *PLoS ONE,* vol. 5, p. 10, 04 05 2015.

[38] D. T. J. C. S. C. S. M. C. D. R. C. M. D. G. L. S. E. Lo Presti D, "Plant Growth Monitoring: Design, Fabrication, and Feasibility Assessment of Wearable Sensors Based on Fiber Bragg Gratings," *Sensors,* vol. 23, no. 1, p. 361, 2023.

[39] B. I. a. R. E. S. F. C. Azhar, "Controlling and monitoring ornamental plants care remotely using android application," in *2017 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob),* Bandung, Indonesia, 2017.

[40] H. H. N. a. A. H. A. A. R. Achmad Fauzy, "Implementation of Monitoring System and Prediction of Room Humidity for Orchid," in *2022 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob),* Bandung, Indonesia, 2022.

[41] U. S. S. Samaratunge Arachchillage, D. H. L. Amarasinghe, M. C. Kirindegamaarachchi, A. B. L. and F. W. M. K. S. S. W, "Smart Intelligent

Floriculture Assistant Agent (SIFAA)," in *2021 3rd International Conference on Advancements in Computing (ICAC)*, Colombo, 2021.

[42] S. Kulasekara, K. Kumarasiri, T. Sirimanna, D. Dissanayake, A. Karunasena and N. Pemadasa, "Machine Learning Based Solution for Improving the Efficiency of Sugar Production in Sri Lanka," in *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kharagpur, 2022.

[43] R. Gamage, H. Rajapaksa, A. Sangeeth, G. Hemachandra, J. Wijekoon and D. Nawinna, "Smart Agriculture Prediction System for Vegetables Grown in Sri Lanka," in *021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, 2021.

[44] Meeradevi, V. Sanjana and M. R. Mundada, "Decision Support System to Agronomically Optimize Crop Yield based on Nitrogen and Phosphorus," in *2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS)*, Bengaluru, 2019.

[45] X. Chen, J. Zhao, J. Bi and L. Li, "Research of real-time agriculture information collection system base on mobile GIS," in *2012 First International Conference on Agro- Geoinformatics (Agro-Geoinformatics)*, Shanghai, 2012.

[46] U. Saini, R. Kumar, V. Jain and M. U. Krishnajith, "Univariant Time Series forecasting of Agriculture load by using LSTM and GRU RNNs," in *2020 IEEE Students Conference on Engineering & Systems (SCES)*, Prayagraj, 2020.

[47] A. Aravatagimath, A. V. Sutagundar and V. Yalavigi, "Agriculture Product Marketing Data Analysis using Machine Learning," in *2021 International Conference on Forensics, Analytics, Big Data, Security (FABS)*, Bengaluru, 2021.

[48] B. V. B. Prabhu and M. Dakshayini, "Demand-prediction model for forecasting AGRI-needs of the society," in *2017 International Conference on Inventive Computing and Informatics (ICICI)*, Coimbatore, 2017.

[49] K. S. Pratyush Reddy, Y. M. Roopa, K. Rajeev L.N. and N. S. Nandan, "IoT based Smart Agriculture using Machine Learning," in *2020 Second*

International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, 2020.

[50] G. Chauhan and A. Chaudhary, "Crop Recommendation System using Machine Learning Algorithms," in *2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART)*, MORADABAD, 2021.

[51] Z. Zhu, R. Zhang and J. Sun, "Research on GIS-Based Agriculture Expert System," in *2009 WRI World Congress on Software Engineering*, Xiamen, 2009.

[52] M. U. A. B. A. R. H. T. E. F. Y. L. J. S. I. H. M. F. M. Matteo G. Ziliani, "Early season prediction of within-field crop yield variability by assimilating CubeSat data into a crop model," 11 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168192321004226. [Accessed 10 03 2023].

[53] B. I. a. R. E. S. R. Ubudi, "Automation system for controlling and monitoring ornamental plants using fuzzy logic method," in *International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC)*, Yogyakarta, Indonesia, 2017.

[54] G. Burchi, S. Chessa, F. Gambineri and A. Kocian, "nformation technology controlled greenhouse: A system architecture," in *IoT Vertical and Topical Summit on Agriculture - Tuscany (IOT Tuscany)*, Tuscany, Italy, 2018.

[55] Y.-S. &. L. T.-H. &. Y. K. S. Tong, "Deep Learning for Image-Based Plant Growth Monitoring: A Review," *International Journal of Engineering and Technology Innovation.*, vol. 12, no. 3, pp. 225-246, 05 2022.

[56] G.-Y. &. L. R.-R. &. C. Z. &. T. K. &. S. M.-F. &. Y. G.-H. &. L. Q.-Q. Zhang, "Manipulation of the Rice L-Galactose Pathway: Evaluation of the Effects of Transgene Overexpression on Ascorbate Accumulation and Abiotic Stress Tolerance," *PLoS ONE*, vol. 10, no. 5, 2015.

[57] D. T. J. C. S. C. S. M. C. D. R. C. M. D. G. L. S. E. Lo Presti D, "Plant Growth Monitoring: Design, Fabrication, and Feasibility Assessment of Wearable Sensors Based on Fiber Bragg Gratings," *Sensors.*, vol. 23, no. 1, p. 361, 2023.

[58] B. I. a. R. E. S. F. C. Azhar, "Controlling and monitoring ornamental plants care remotely using android application," in *2017 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, Bandung, Indonesia,, 2017.

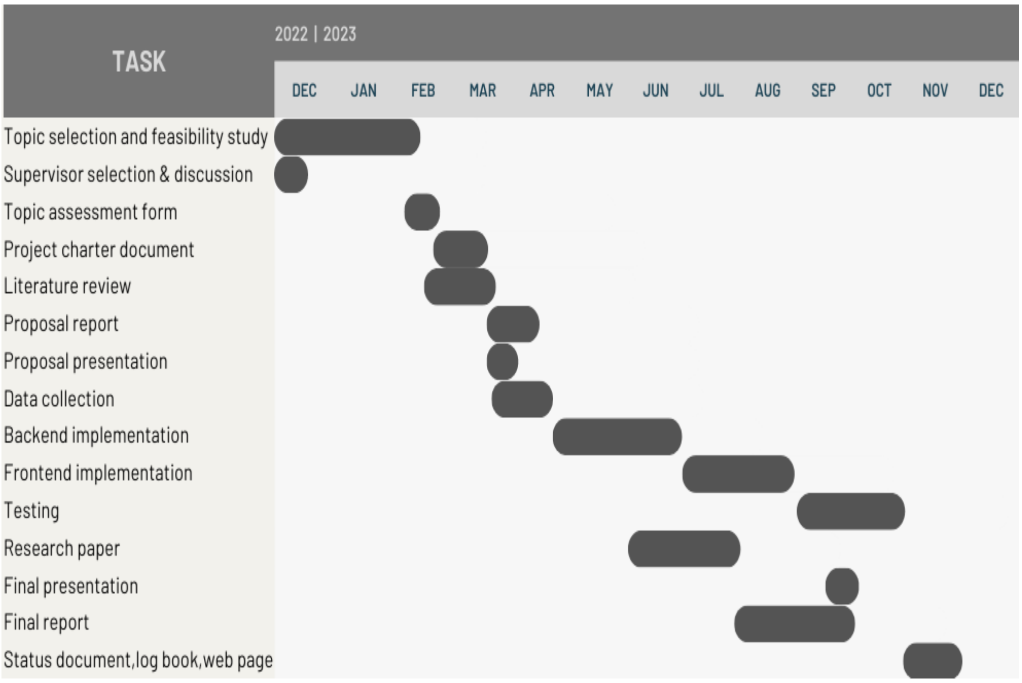
[59] H. H. N. a. A. H. A. A. R. Achmad Fauzy, "Implementation of Monitoring System and Prediction of Room Humidity for Orchid," in *2022 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, Bandung, Indonesia, 2022.

APPENDICES

Appendix 1 : Survey on Floriculture Industry

https://forms.gle/mknPpztYp63e2wJE8

Appendix 2 : Gantt Chart



Appendix 3 : 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