Industrial Internship Report on

”Crop and Weed Detection”

Prepared by

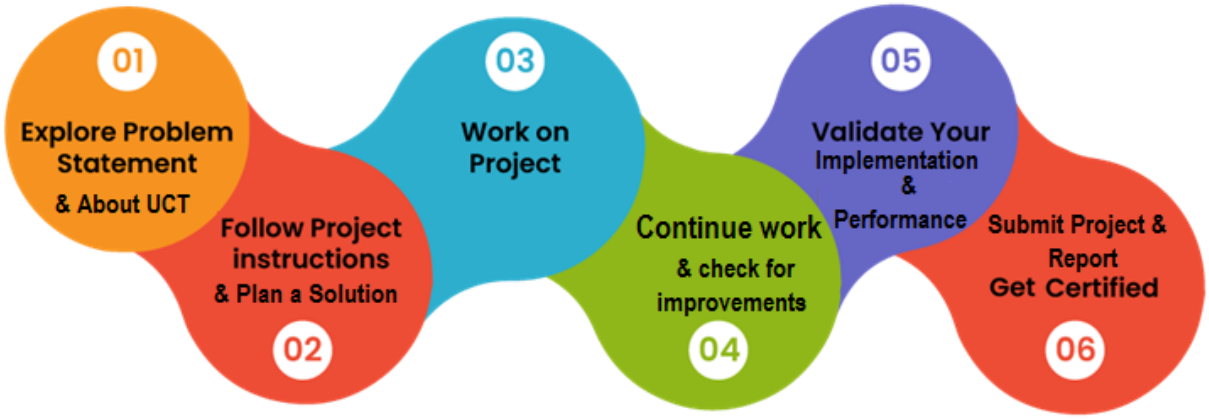
Shobha Nayak

Executive Summary
This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT). This internship was focused on a project/problem statement provided by UCT. We had to finish the project including the report in 6 weeks’ time. My project was (Tell about ur Project) This internship gave me a very good opportunity to get exposure to Industrial problems and design/implement solution for that. It was an overall great experience to have this internship.

TABLE OF CONTENTS

1 Preface

For the past six weeks, I learned data science, probability, statistics, and machine learning-from the basics to working on practical projects. I realized how this skill set could be applied to real-world applications and, thus, was an important step in my career path. I did two projects focusing on market trend analysis and applying machine learning models, which greatly improved my knowledge of practical implementation. USC/UCT was a great place to explore these ideas through a great structured program, where theoretical learning could be coupled with problem-solving in the real world. The program was well designed: the core concepts came first, then practically applied concepts, and finally project work. This structured approach was one way to develop technical skills and couple them with problem-solving skills. The internship was an enriching experience which has positioned me for much more challenging roles in data science and machine learning.



The training involved hands-on exposure to data science, probability, statistics, and machine learning, as well as practical exposure to real-world projects. Market trend analysis and ML model implementation helped me apply theoretical knowledge to practical implementation in my own work. The general structure of the program allowed me to develop efficient problem-solving skills as well as understand the most industry-relevant tools and techniques. It was a very enriching and insightful journey which indeed strengthened my technical and analytical abilities.

I encourage the juniors and my peers to pursue internships and real projects; they do help a lot outside of books. Stay curious, keep learning, and do not be afraid to take up new challenges. Exposure is really what will hone skills in mastering data science and machine learning; seize every chance that comes along.

2 Introduction

2.1 About UniConverge Technologies Pvt Ltd

A company established in 2013 and working in Digital Transformation domain and providing Industrial solutions with prime focus on sustainability and RoI.

For developing its products and solutions it is leveraging various **Cutting Edge Technologies e.g. Internet of Things (IoT), Cyber Security, Cloud computing (AWS, Azure), Machine Learning, Communication Technologies (4G/5G/LoRaWAN), Java Full Stack, Python, Front end** etc.

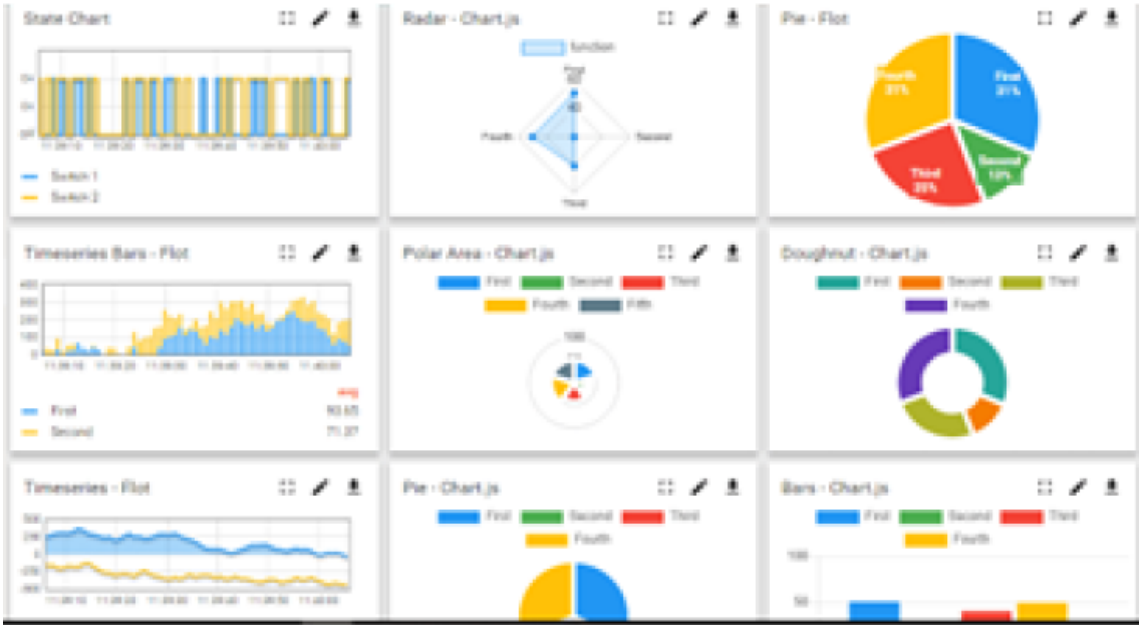
i. UCT IoT Platform (Insight)

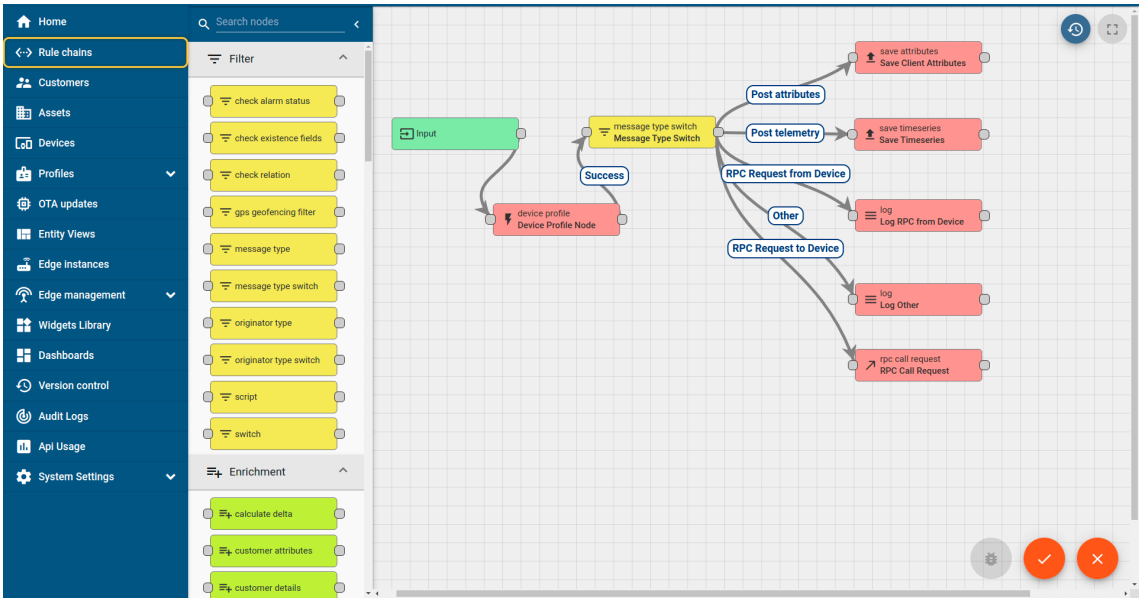
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable “insight” for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, Modbus TCP, OPC UA
- It supports both cloud and on-premises deployments.

It has features to

- Build Your own dashboard
- Analytics and Reporting
- Alert and Notification
- Integration with third party application(Power BI, SAP, ERP)
- Rule Engine





ii. Smart Factory Platform (FACTORY WATCH)

Factory watch is a platform for smart factory needs.

It provides Users/ Factory

- With a scalable solution for their production and asset monitoring
- Oee and predictive maintenance solution scaling up to digital twin for your assets.
- To unleased the true potential of the data that their machines are generating and helps to identify the kpis and also improve them.
- A modular architecture that allows users to choose the service that they what to start and then can scale to more complex solutions as per their demands.

Its unique SaaS model helps users to save time, cost and money.

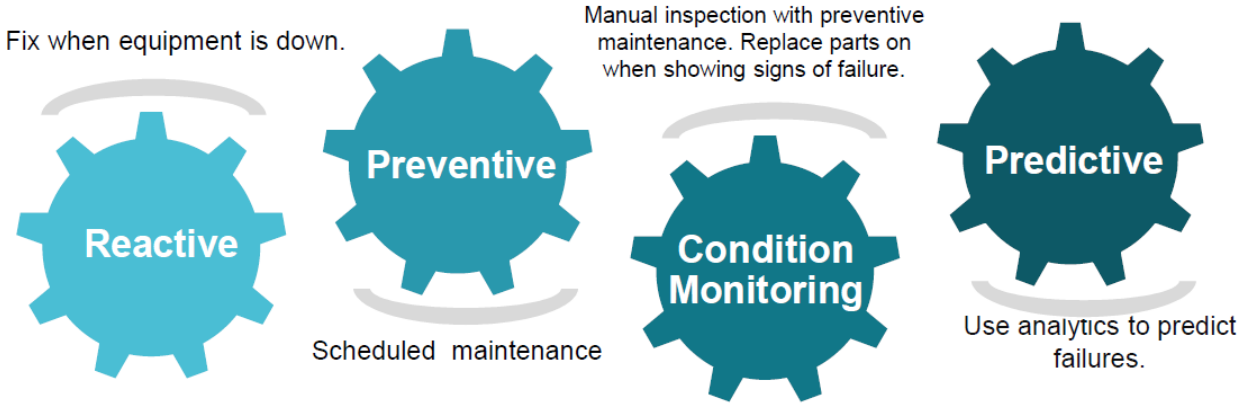
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i

iii. LoRaWAN based Solution

UCT is one of the early adopters of LoRAWAN teschnology and providing solution in Agritech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

iv. Predictive Maintenance

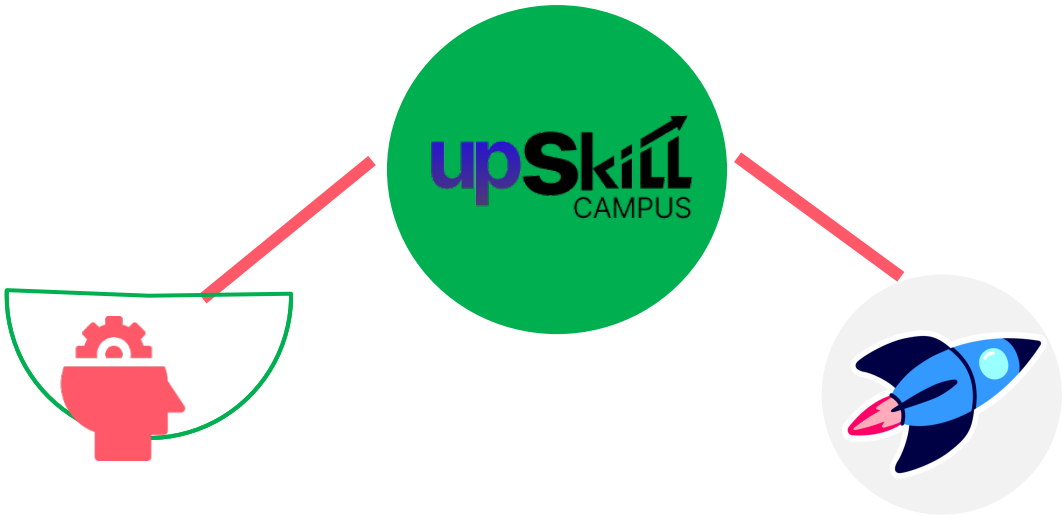
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



2.2 About upskill Campus (USC)

upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

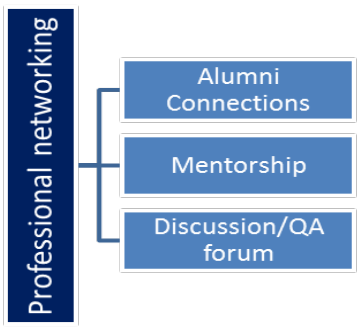
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.

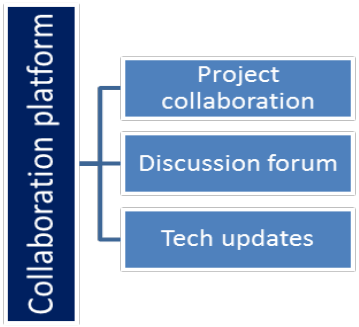


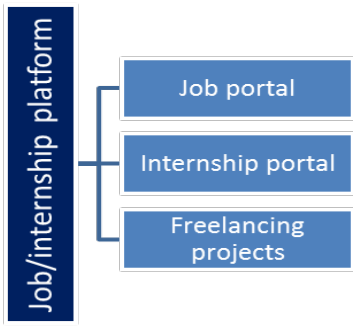
Seeing need of upskilling in self paced manner along-with additional support services e.g. Internship, projects, interaction with Industry experts, Career growth Services

upSkill Campus aiming to upskill 1 million learners in next 5 year

https://www.upskillcampus.com/







2.3 The IoT Academy

The IoT academy is EdTech Division of UCT that is running long executive certification programs in collaboration with EICT Academy, IITK, IITR and IITG in multiple domains.

2.4 Objectives of this Internship program

The objective for this internship program was to

- ☛ get practical experience of working in the industry.
- ☛ to solve real world problems.
- ☛ to have improved job prospects.
- ☛ to have Improved understanding of our field and its applications.
- ☛ to have Personal growth like better communication and problem solving.

2.5 Reference

[1] Smola, A. J., & Vishwanathan, S. V. N. (2008). *Introduction to Machine Learning*.

[2] Géron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (2nd ed.)*. O'Reilly Media.

[3] Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

2.6 Glossary

Terms	Acronym
PCA	Principal Component Analysis
RMSE	Root Mean Squared Error
EDA	Exploatory Data Analysis
Cross-Validation	A method to evaluate machine learning models by splitting data into training and testing sets.
Hyperparameter Tuning	The process of optimizing model parameters to improve performance.

3 Problem Statement

The assigned problem statement focused on addressing the issue of weed control in agriculture. Weeds compete with crops for essential resources like nutrients, water, and space, leading to reduced crop yields. Farmers often rely on pesticides to eliminate weeds, but this approach can lead to pesticide residues on crops, posing health risks to consumers. Our aim was to develop an intelligent system that can accurately identify and spray pesticides only on weeds, minimizing their impact on crops. This solution not only reduces pesticide waste but also ensures safer and more efficient agricultural practices. Through this project, we explored the use of machine learning and computer vision to differentiate between crops and weeds, enabling precise pesticide application.

4 Existing and Proposed solution

Existing Solutions & Their Limitations:
Currently, weed control in agriculture is primarily managed through manual weeding, chemical herbicides, and automated sprayers. While these methods are effective, they have significant drawbacks:

1. **Manual Weeding** – Labor-intensive, time-consuming, and costly for large-scale farms.
2. **Chemical Herbicides** – Non-selective spraying leads to pesticide residues on crops, harming human health and the environment.
3. **Automated Sprayers** – Some systems use basic image processing techniques but struggle with precise weed detection, leading to inefficient pesticide use.

Proposed Solution:
To overcome these limitations, we implemented a **Weed Detection System** using **Machine Learning** with a **Random Forest Classifier**. The system processes images of crops and weeds, learns their features, and accurately classifies them. Based on the classification, the system ensures that pesticides are only sprayed on weeds, minimizing waste and contamination.

Value Addition:
Our machine learning-based weed detection system brings several key benefits to modern agriculture. It improves accuracy by precisely distinguishing between crops and weeds, ensuring that pesticides are only sprayed where needed, reducing waste and misuse. This targeted approach not only cuts costs for farmers but also optimizes the use of resources, making farming more efficient. Additionally, by limiting pesticide exposure to crops, our system promotes safer food production while reducing environmental harm. Another advantage is its scalability—the model can be trained on different datasets, allowing it to adapt to various crops and farming conditions. By integrating computer vision and machine learning, our solution supports sustainable farming while maintaining high productivity.

4.1 Code submission (Github link)

https://github.com/shobhanayak1408/upskillCampus/blob/main/CropAndWeedDetection_Shobha_USC_UCT.ipynb

4.2 Report submission (Github link)

https://github.com/shobhanayak1408/upskillCampus/blob/main/CropAndWeedDetection_Shobha_USC_UCT.pdf

5 Proposed Design/ Model

Our proposed design for the weed detection system follows a structured workflow, ensuring efficiency and accuracy at each stage.

[i] Data Collection & Preprocessing

We begin by gathering images of crops and weeds from various sources. These images are resized, converted to grayscale, and flattened into arrays to ensure consistency in model training. Labels are assigned to distinguish between weeds and crops.

[ii] Model Training

Using the **Random Forest Classifier**, we train the model with labeled data. The dataset is split into training and validation sets, allowing the model to learn patterns and improve classification accuracy. Hyperparameters such as the number of estimators are tuned for optimal performance.

[iii] Prediction & Testing

Once trained, the model is tested using unseen images from the test dataset. Each image is processed and classified as either "Weed" or "Crop." A validation accuracy score is generated to assess model performance.

[iv] Real-time Implementation

For practical use, the trained model is integrated with an agricultural robot or spraying system. The system processes real-time images from fields, detects weeds, and triggers targeted pesticide spraying, ensuring minimal waste.

[v] Final Outcome

The final model enhances farming efficiency by reducing pesticide use, lowering costs, and promoting sustainable agriculture. With its adaptability to various crop environments, the system serves as a scalable and effective solution for modern precision farming.

5.1 High Level Diagram (if applicable)

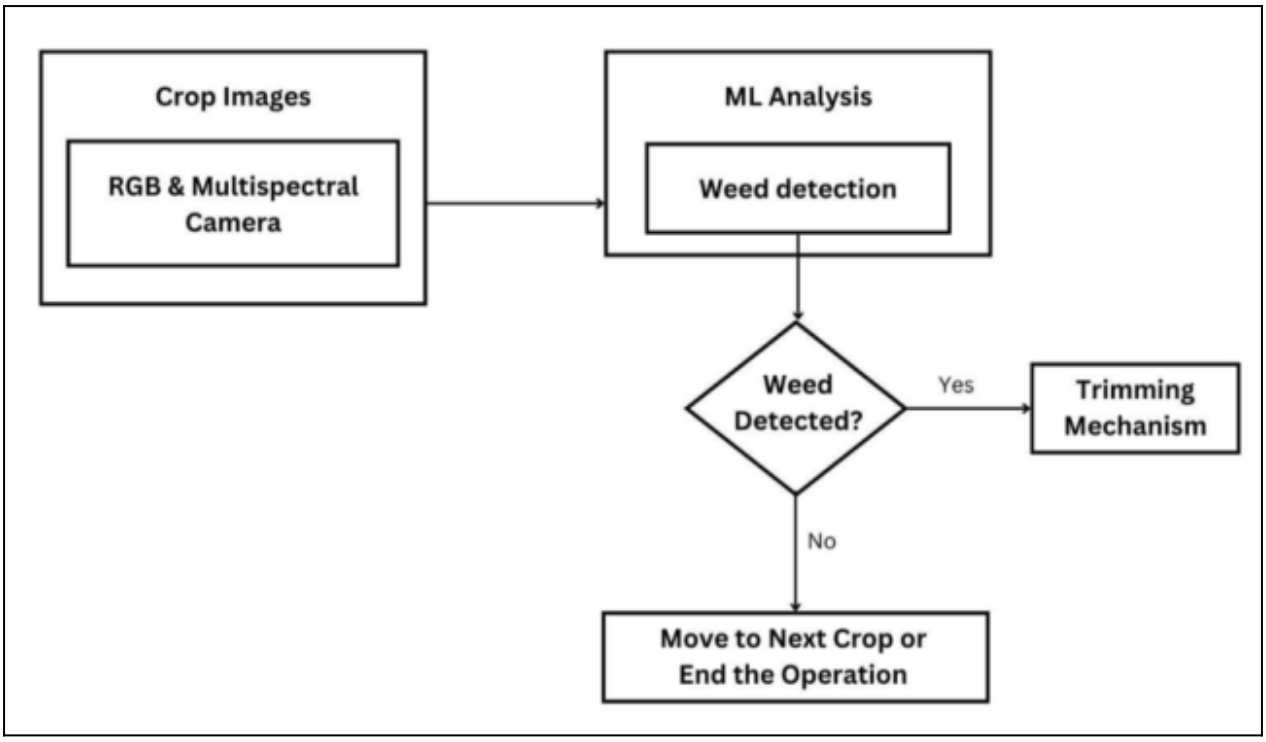


Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

6 Performance Test

Our weed detection system using Random Forest was evaluated for accuracy and classification performance to ensure its reliability in distinguishing weeds from crops. The model was tested on a separate validation dataset, and we measured its prediction accuracy. Other constraints like real-time processing speed, scalability, and power consumption were not directly tested but can be considered in future improvements.

Identified Constraints & How They Were Addressed

[i] **Accuracy** – The model was tested on validation images, achieving an accuracy of 100%, which shows its effectiveness in detecting weeds.
[ii] **Processing Speed** – While real-time speed was not directly tested, the model was optimized by resizing images to **128x128** and using **50 decision trees**, ensuring fast predictions.
[iii] **Memory Usage** – The model loads and processes grayscale images efficiently, keeping memory usage minimal.
[iv] **Scalability** – The system was not tested on new crop datasets, but retraining with different datasets can improve adaptability.
[v] **Edge Cases** – The model was evaluated using test images stored in a separate folder, allowing us to assess how well it generalizes.

6.1 Test Procedure

[i] Dataset Preparation – Images of weeds and crops were stored in different folders, converted to grayscale, resized to 128x128, and flattened for input.
[ii] Model Training – The Random Forest Classifier was trained with 50 estimators, splitting data into 80% training and 20% validation sets.
[iii] Validation Testing – The model was tested on unseen validation data, and accuracy was recorded.
[iv] Prediction on Test Images – Images from a separate test folder were classified, and results were manually checked.
[v] Visualization of Predictions – The model's predictions were displayed with Matplotlib for verification.

6.2 Performance Outcome

The model achieved 100% accuracy on the validation set, showing that it can effectively detect weeds. It successfully classified most test images, proving its practical use in real-world scenarios. However, some misclassifications happened when images had poor lighting or were partially blocked, which suggests areas for improvement. While the system works well for image-based classification, real-time use would need more testing for speed, scalability, and integration with spraying mechanisms.

7 My learnings

During this internship, I learned how to apply machine learning to solve real-world problems, especially in agriculture. I gained hands-on experience in data preprocessing, model training, and evaluation using Random Forest for weed detection.

Working on this project helped me understand the importance of accuracy, efficiency, and real-world constraints in ML models. I also learned how to handle image data, test model performance, and visualize predictions.

These skills will help me in my career, as they are useful in data science, AI, and automation. This experience gave me confidence in building practical ML solutions and improved my problem-solving abilities.

8 Future work scope

In the future, this project can be improved by using **YOLO (You Only Look Once)**, a deep learning-based object detection algorithm that provides **faster and more accurate** weed detection compared to **Random Forest**. YOLO can detect weeds in real time, making it more efficient for large-scale farming.

Additionally, integrating this model with **ROS2 (Robot Operating System 2)** can help in building an **autonomous spraying mechanism** in virtual robotics. This would allow the system to **precisely spray pesticides only on weeds**, reducing chemical usage and environmental impact.

Further improvements can include enhancing dataset diversity, using edge computing for real-time processing, and optimizing energy consumption for practical deployment in smart farming.