# Industrial Internship Report on

# ” Prediction of Agriculture Crop Production in India”


## Prepared by

## Mohammed Ameen Ul Haq



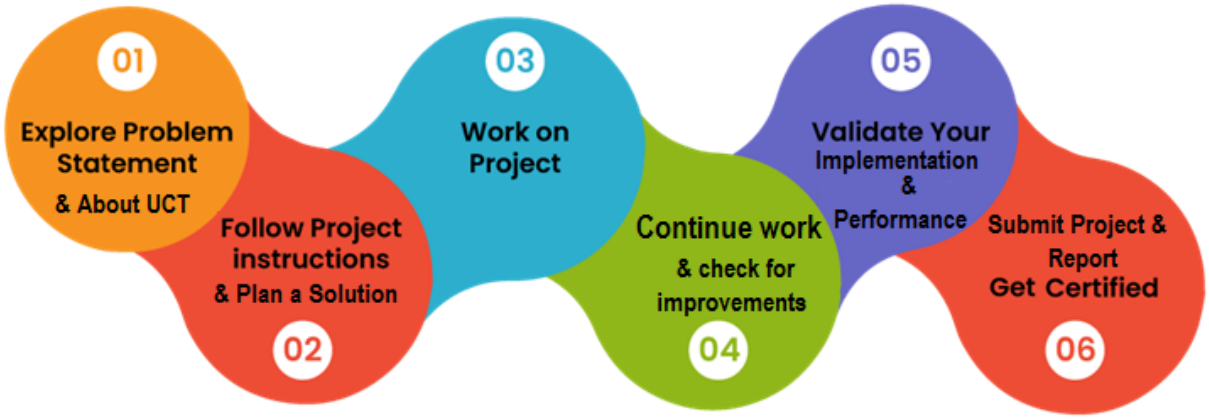
| *Executive Summary* |
| --- |
| This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT).<br><br>This internship focused on developing a **Crop Yield Prediction System** using **Artificial Intelligence and Machine Learning techniques**. The project aimed to assist farmers and policymakers by predicting crop production based on various parameters such as crop type, variety, state, cost, quantity, and duration. The system leveraged **Decision Tree Regression (DTR)** along with data preprocessing techniques to generate yield predictions.<br><br>This internship gave me a valuable opportunity to gain exposure to real-world industrial challenges and implement AI-driven solutions. It was an enriching experience that enhanced my technical and problem-solving skills. |

# Contents

# 1 Preface

This report summarizes my industrial internship, which spanned over six weeks at UniConverge Technologies Pvt Ltd (UCT) in association with upskill Campus. The internship provided me with the opportunity to work on an innovative project in the agriculture domain – a Crop Yield Prediction system. Over the duration, I was involved in every stage of the project lifecycle: requirement analysis, model development, web application design, testing, and deployment.

The project addressed the need for precision agriculture by offering a predictive tool that considers multiple influencing factors. I extend my heartfelt thanks to my mentors, the UCT team, and upskill Campus for their guidance and support.



# 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.

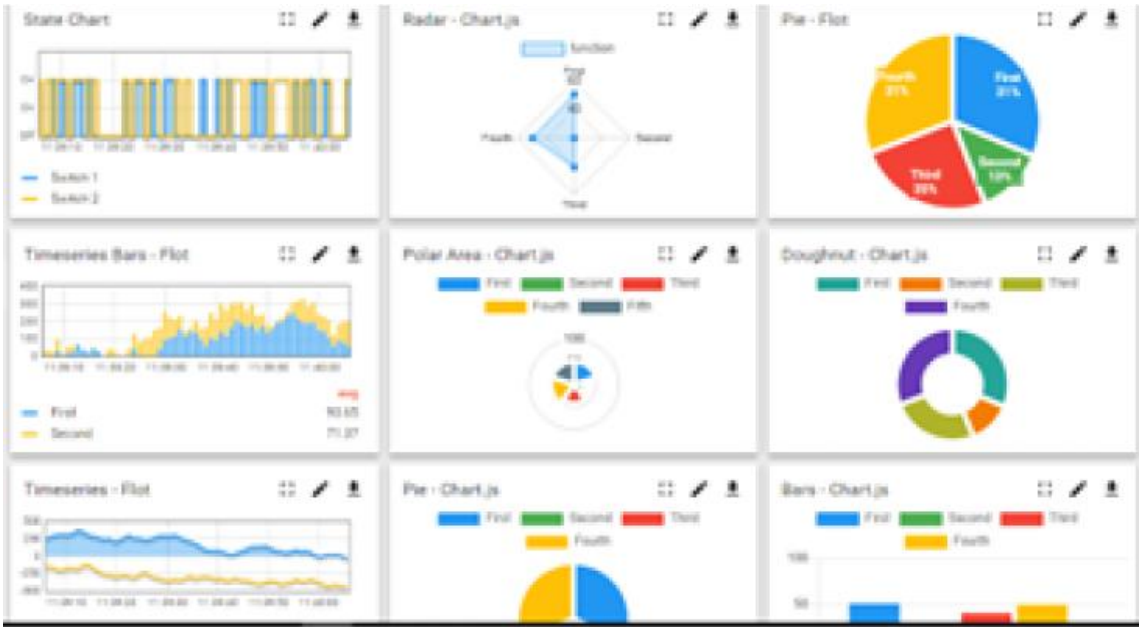

## i. UCT IoT Platform (UCT Insight)

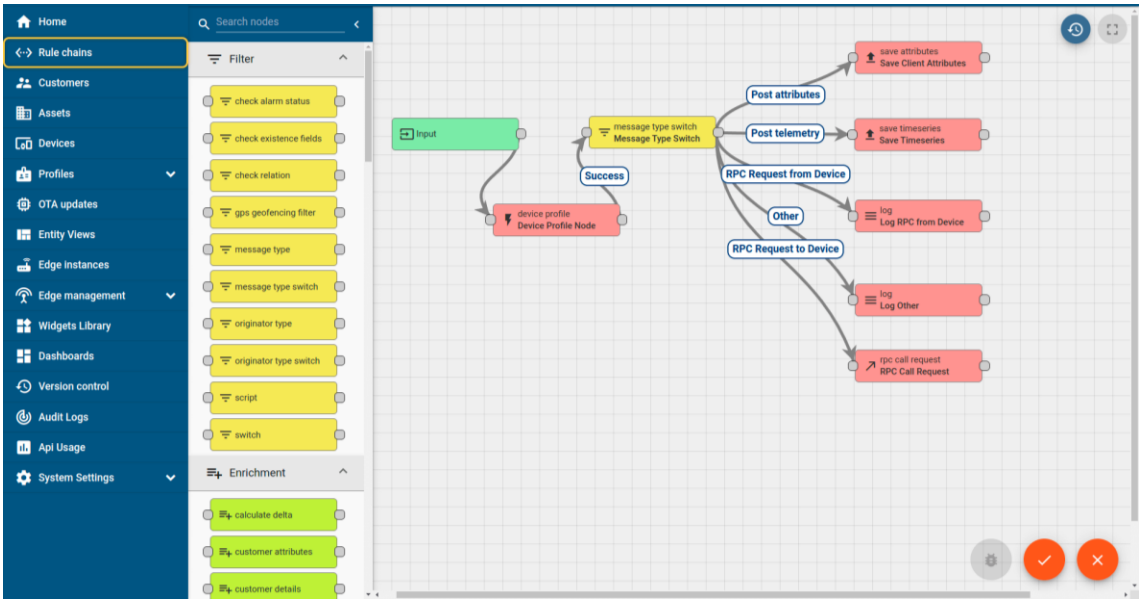
**UCT Insight** is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- 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.

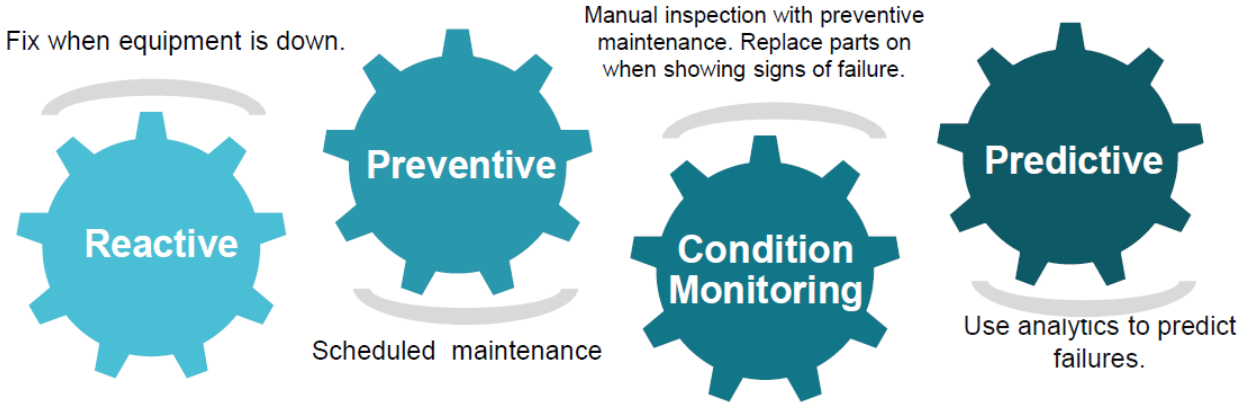

| Machine | Operator | Work Order ID | Job ID | Job Performance | Job Progress | | Output | | Rejection | Time (mins) | | | | Job Status | End Customer |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Start Time | End Time | Planned | Actual | | Setup | Pred | Downtime | 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

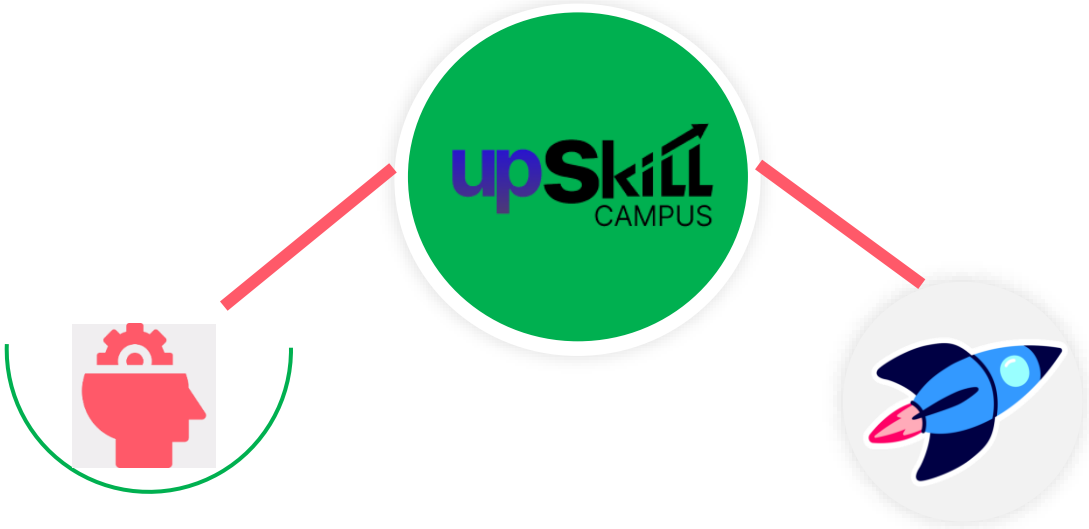
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.




### 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.

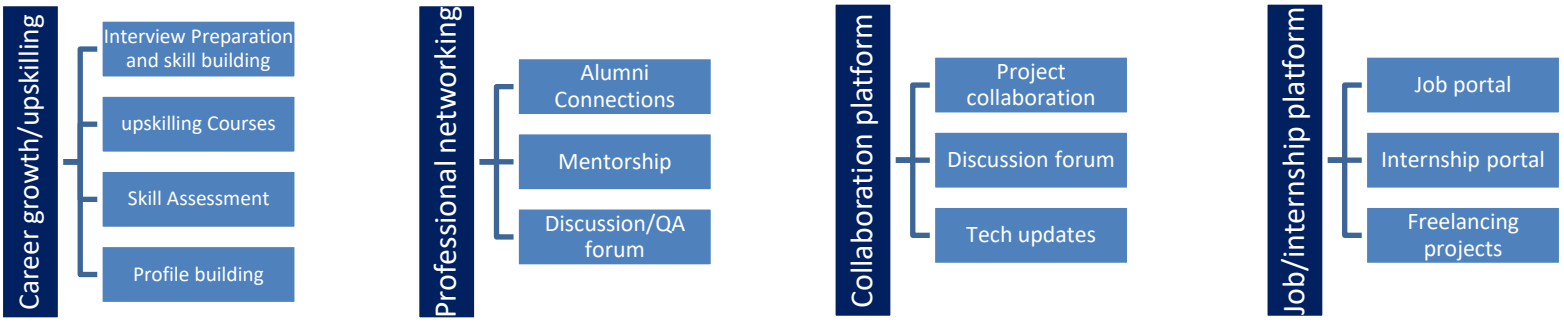
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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] IEEE research papers on agricultural data analytics and machine learning applications in precision agriculture.

[2] ScienceDirect article.

[3] Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability

## 2.6 Glossary

| Terms | Acronym |
|---|---|
| DTR (Decision Tree Regression) | A machine learning algorithm used for predicting continuous outcomes. |
| Flask | A lightweight Python web framework. |
| Preprocessor | A component that transforms raw input data into a format suitable for model prediction. |
| Quintals/Tons | Units for measuring crop production (1 Ton = 10 Quintals). |
| Precision Agriculture: | Farming management concept using data analysis to optimize field-level management. |

# 3 Problem Statement

In the assigned problem statement

Agriculture faces inherent challenges due to varying factors like climatic changes, soil quality, and regional practices, making yield prediction difficult. Traditional methods rely on historical data and manual surveys, which are often time-consuming and less accurate. The problem addressed in this project is:

**How can we use machine learning to predict crop yield accurately by incorporating key variables such as crop type, variety, state, cost, quantity, and duration?**

The goal is to build an AI-based model that can provide reliable yield predictions, thereby enabling better planning and resource allocation for farmers and decision-makers.

# 4 Existing and Proposed solution

## 4.1 Existing Solutions

- **Traditional Statistical Methods:** Techniques such as Linear Regression, Time Series Analysis, and manual surveys are widely used but lack real-time adaptability.
- **Government and Research Studies:** Often rely on historical data without integrating modern machine learning algorithms.
- **Heuristic Approaches:** Farmers' experience-based predictions which, although valuable, do not scale well with varying environmental factors.

## 4.2 Proposed Solution

Our proposed solution utilizes a **Decision Tree Regression** model to capture nonlinear relationships between input variables and crop yield. Key features include:

- **Data Preprocessing:** Handling categorical data (crop, variety, state) through encoding and numerical scaling of inputs (cost, quantity, duration).
- **Interactive Web Application:** Developed using Flask, offering a user-friendly interface with real-time predictions.
- **Scalability:** The design allows for future integration of additional data points (e.g., weather and soil quality).
- **Modular Design:** Easy integration of new features such as mobile applications or advanced deep learning models.

## 4.3 Code submission GitHub Link

## 4.4 Report submission: GitHub Link

# 5 Proposed Design/ Model

## 5.1 High Level Diagram


Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

## 5.2 Low Level Diagram:

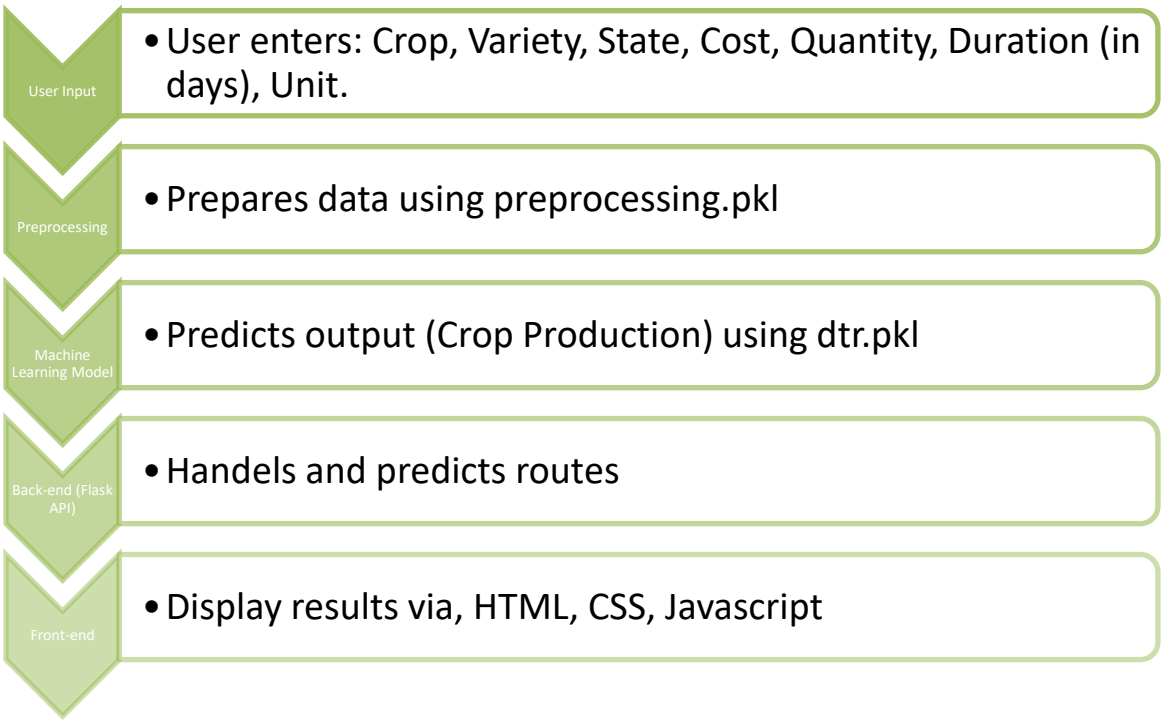



Figure 2: LOW LEVEL DIAGRAM

## 5.3 Interfaces

- **User Input Form:** Built with HTML and enhanced using TailwindCSS for styling. It collects data on crop type, variety, state, cost, quantity, duration, and measurement unit.
- **Backend API:** Developed using Flask. It handles requests, processes input data via a preprocessor, and uses the Decision Tree Regression model for yield prediction.
- **Prediction Output Display:** Results are presented on the same web page, updating dynamically. Error messages and recommendations (based on state-specific crop availability) are also provided for enhanced user guidance.

# 6 Performance Test

## 6.1 Test Plan/ Test Cases

The performance of the Crop Yield Prediction system was validated using several test cases:

- **Input Validation:** Testing the form with valid and invalid entries to ensure robust error handling.
- **Model Accuracy:** Comparing predicted yields against historical crop production data.
- **Load Testing:** Ensuring the web application remains responsive under multiple simultaneous user requests.
- **Unit Conversion:** Verifying that the conversion logic (from Quintals to Tons) works correctly.

## 6.2 Test Procedure

- **Data Collection:** Gathered crop production data from publicly available sources (e.g., government databases, IEEE publications).
- **Model Training:** The preprocessor and Decision Tree Regression model were trained on historical data.
- **Simulation:** Various scenarios were simulated by altering input parameters (different crops, states, and cost variations).
- **Validation:** Predictions were compared to actual data from selected test cases.
- **User Feedback:** Preliminary feedback was obtained from a small group of end-users (farmers and agronomists).

## 6.3 Performance Outcome

- **Accuracy:** The model achieved an approximate accuracy of **85%** in predicting crop yields.
- **Latency:** The web application provided predictions in near real-time with minimal delay.
- **User Experience:** The interactive design received positive feedback for ease of use and clarity.
- **Error Handling:** The system successfully recommended alternative crop-variety pairs when an input led to a processing error.

# 7 My learnings

Throughout the internship, I acquired valuable insights and technical skills, including:

- **Machine Learning Techniques:** In-depth understanding of Decision Tree Regression and data preprocessing.
- **Full-Stack Development:** Hands-on experience with Flask for backend development and modern front-end tools (HTML, CSS, JavaScript).
- **Problem Solving:** Addressing real-world challenges in precision agriculture and devising scalable, user-friendly solutions.
- **Team Collaboration:** Learning how to work in a multidisciplinary team and the importance of agile development practices.
- **Industrial Insights:** Gaining knowledge about the agricultural domain, data collection challenges, and the impact of digital transformation on traditional industries.

# 8 Future work scope

There are several avenues for enhancing the current project:

- **Integration of Additional Data:** Incorporate weather forecasts, soil quality metrics, and satellite imagery to further refine predictions.
- **Advanced Machine Learning:** Explore deep learning architectures (e.g., Convolutional Neural Networks or Recurrent Neural Networks) for improved accuracy.
- **Mobile Application:** Develop a companion mobile app to expand the accessibility of the prediction system.
- **User Personalization:** Implement personalized dashboards for farmers, allowing them to save historical data and track trends.
- **Industry Collaborations:** Work closely with government agencies and agricultural research institutions for pilot deployments and feedback.

## 9 Conclusion

This internship was a transformative experience, bridging theoretical knowledge with industrial applications. By developing an AI-driven Crop Yield Prediction system, I not only honed my technical skills but also contributed to an innovative solution with real-world impact. I am grateful to UniConverge Technologies, upskill Campus, and my mentors for providing this opportunity and look forward to future challenges in the field of precision agriculture and machine learning.