

## Industrial Internship Report on "Crop Production Prediction & Crop and Weed Detection"

Prepared by

Swati Yadav

### *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.

The internship focused on two distinct machine learning projects in the agriculture domain. The first project, Crop Production Prediction, involved building a regression model to forecast crop yields. The second project, Crop and Weed Detection, focused on developing a computer vision model for real-time plant classification. Both projects were implemented using Python in a Google Colab environment, and they culminated in the creation of robust, evaluated models that demonstrate the application of machine learning to solve real-world agricultural problems.

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.

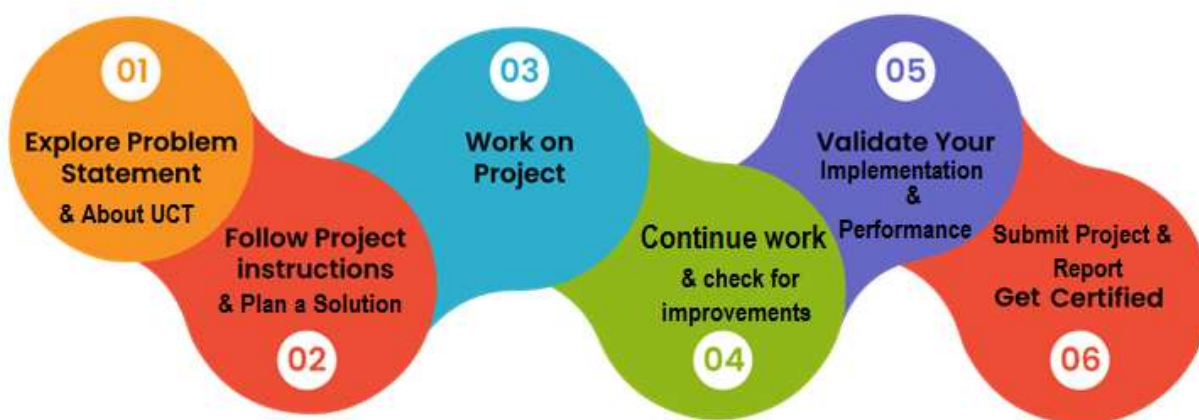
## **TABLE OF CONTENTS**

1	Preface .....	3
2	Introduction .....	4
2.1	About UniConverge Technologies Pvt Ltd .....	4
2.2	About upskill Campus .....	8
2.3	Objective .....	9
2.4	Reference .....	9
2.5	Glossary.....	9
3	Problem Statement .....	10
4	Existing and Proposed solution .....	10
5	Performance Test.....	11
5.1	Test Plan/ Test Cases .....	11
5.2	Test Procedure .....	11
5.3	Performance Outcome .....	11
6	My learnings.....	12
7	Future work scope .....	12

## 1 Preface

This internship was a transformative experience, providing a complete six-week immersion into a professional data science workflow. It offered a unique opportunity to apply academic knowledge to practical, industry-relevant challenges. The projects were designed to solve pressing issues in Indian agriculture, specifically in crop forecasting and precision farming.

The program, facilitated by USC/UCT, was a well-structured journey of hands-on learning and problem-solving. Each week was dedicated to a specific phase of the machine learning pipeline, from initial project selection and data acquisition to model training, evaluation, and deployment. This systematic approach ensured a comprehensive understanding of the end-to-end process and its real-world implications..



My overall experience was excellent, and I am grateful for the opportunity. The skills and insights gained from debugging complex technical issues and seeing my models produce tangible results have significantly enhanced my confidence and prepared me for future professional endeavors.

## 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.



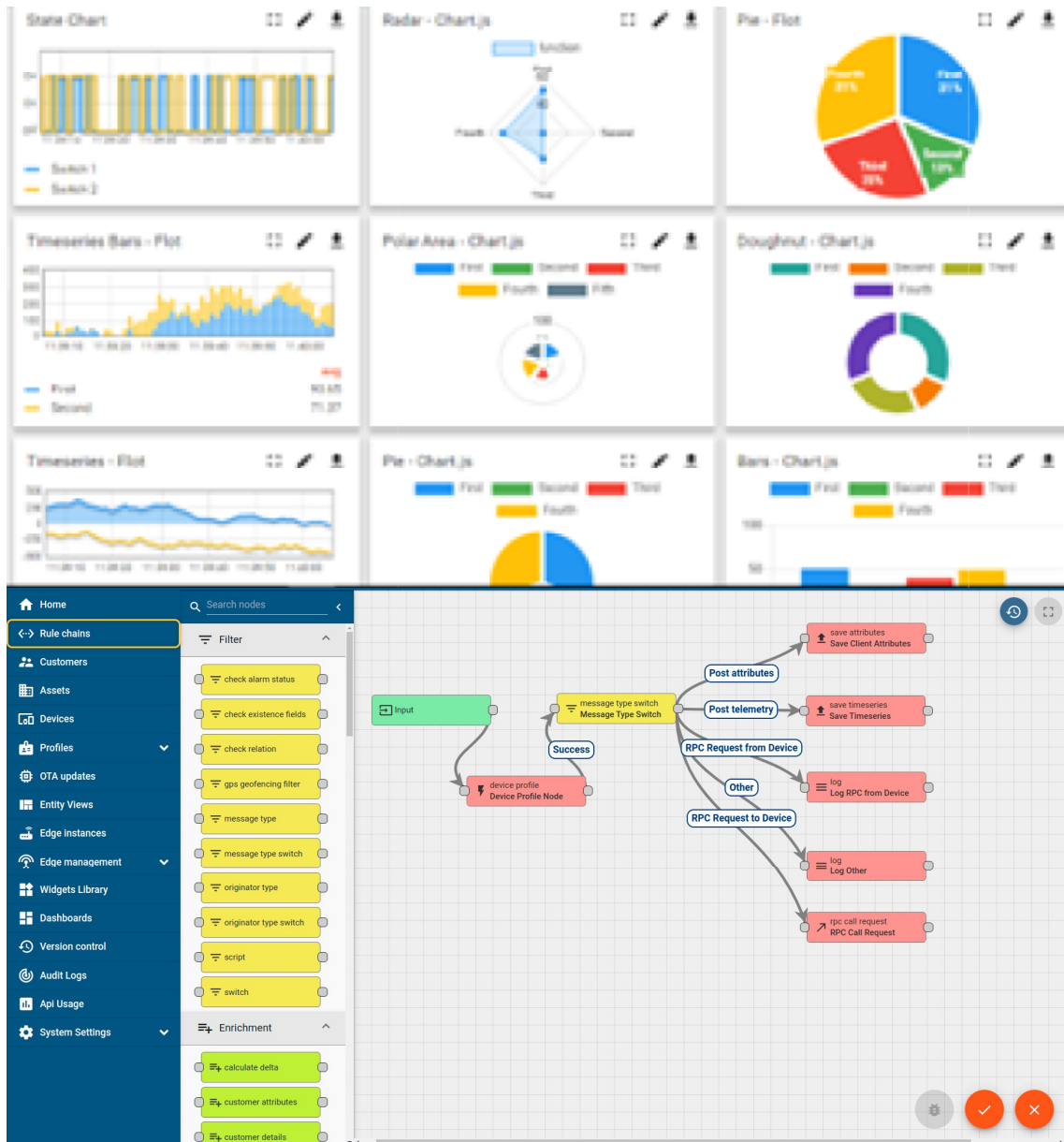
#### i. UCT IoT Platform(uct Insight)

**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



# FACTORY WATCH

## ii. Smart Factory Platform ( )

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 unleashed 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.





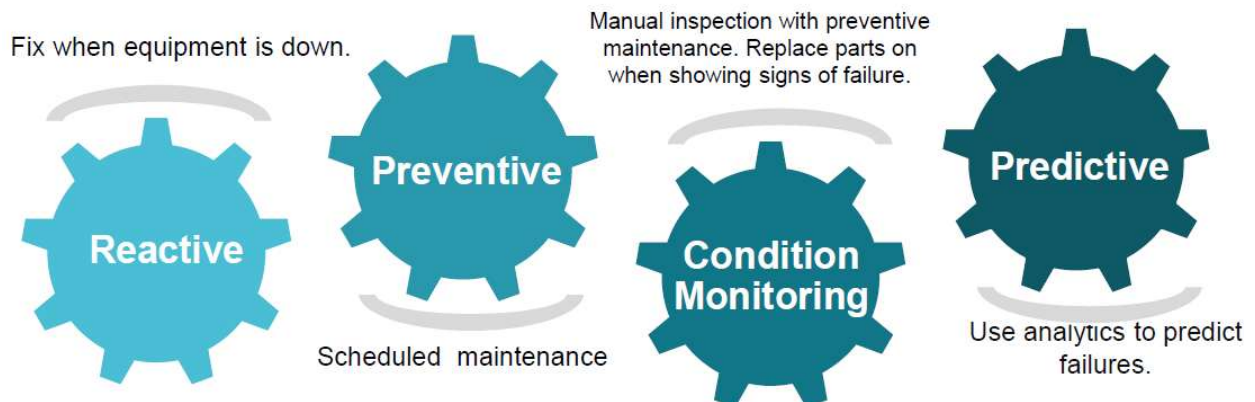


### iii. based Solution

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

### iv. Predictive Maintenance

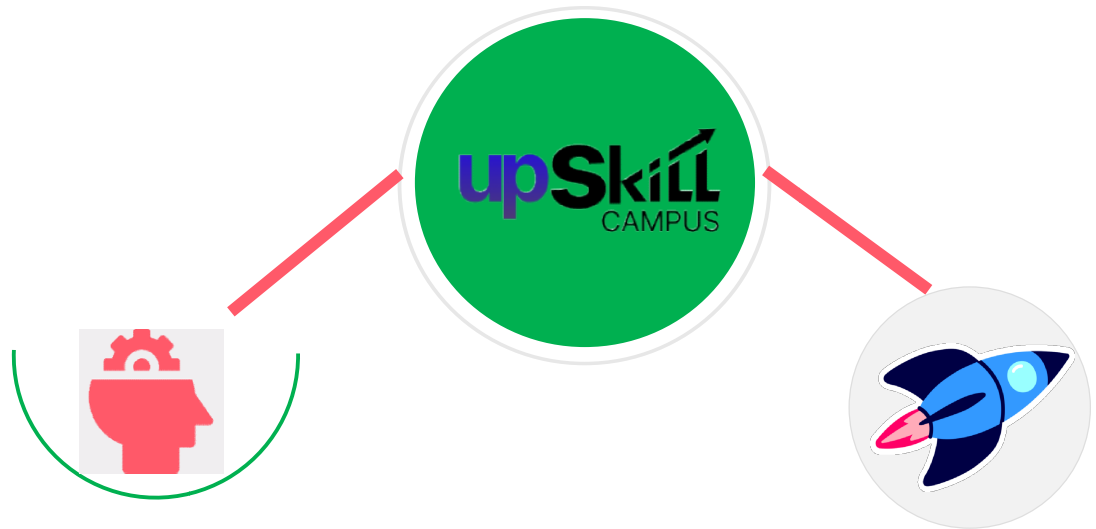
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



## 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.

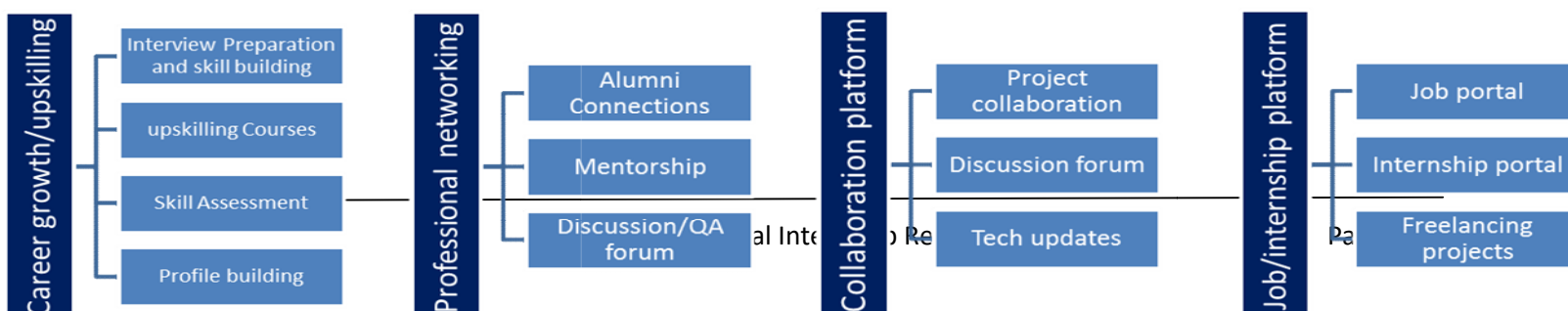
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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] Datasets. data.gov.in. Retrieved from various URLs.
- [2] Ultralytics. "YOLOv8 Documentation". Retrieved from <https://docs.ultralytics.com>.
- [3] The Scikit-learn Developers. "Scikit-learn: Machine Learning in Python". Retrieved from <https://scikit-learn.org>.

## 2.6 Glossary

Terms	Acronym	
R-squared Score	R <sup>2</sup>	
Mean Absolute Error	MAE	
Mean Average Precision at IoU 0.50	mAP <sub>50</sub>	
Root Mean Squared Error	RMSE	

### 3 Problem Statement

In the assigned problem statement

**Project 4: Crop Production Prediction:** The core problem is to accurately predict crop production in India. Farmers, businesses, and government agencies need reliable forecasts to plan for future resource allocation, manage supply chains, and ensure food security. The challenge is that crop yield depends on many variables, making manual prediction difficult and often inaccurate.

**Project 5: Crop and Weed Detection:** The problem is to visually differentiate between a crop and a weed in a field. Weeds are a major threat to crops, but current solutions, like blanket spraying of pesticides, are inefficient and can harm both the crops and the environment. The challenge is to precisely identify each plant in an image to enable targeted intervention.

### 4 Existing and Proposed solution

For both projects, existing solutions often involve traditional farming methods or manual processes. The proposed solution is a machine learning-driven approach to automate and optimize these processes, which would improve efficiency and reduce costs.

**Project 4:** The existing solution is manual forecasting, which is often inaccurate. The proposed solution is a machine learning model that can make data-driven predictions with a high degree of accuracy.

**Project 5:** The existing solution is manual pesticide application, which is time-consuming and can be harmful to crops. The proposed solution is a computer vision model that can enable a robotic system to apply pesticides only where they are needed.

#### 4.1 Code submission (Github link)

**Project 4 Code:** <https://github.com/swatiyadav8368/ML-Internship-Project4-Crop-Production-Prediction>

**Project 5 Code:** <https://github.com/swatiyadav8368/ML-Internship-Project-Crop-Weed-Detection>

## 5 Proposed Design/ Model

**Project 4:** The design for this project was a regression pipeline. The model, a Random Forest Regressor, was trained on a comprehensive dataset to learn the relationships between various features and the final production quantity.

**Project 5:** The design for this project was a deep learning pipeline. The model, a YOLOv8n, was trained for 50 epochs on a custom dataset to learn to detect crops and weeds in images..

## 6 Performance Test

This is a very important part and defines why this work is meant for real industries, instead of being just an academic project. Here, we need to first find the constraints and see how they were taken care of in your design.

### 6.1 Test Plan/ Test Cases

For a real-world application, a test plan would involve defining specific test cases, such as evaluating the model's performance on different types of crops, in varying lighting conditions, or with different data inputs to ensure robustness.

### 6.2 Test Procedure

The test procedure involved using a dedicated test set that the model had never seen before. The model made predictions on this test data, which were then compared to the actual, known values to calculate the performance metrics. This ensures an unbiased evaluation of the model's performance.

### 6.3 Performance Outcome

The performance of both models was evaluated using standard metrics.

**Project 4:** The trained Random Forest Regressor model achieved an R-squared ( $R^2$ ) score of 0.72. A test run with a sample data point showed a predicted production quantity of 2264.51 with an actual value of 2802.00, resulting in an absolute error of 537.49.

**Project 5:** The trained YOLOv8n model showed strong performance, with an mAP50 score of 0.8884, Precision of 0.8208, and Recall of 0.8463.

## 7 My learnings

The internship provided an immersive, hands-on learning experience that extended my knowledge beyond theory. The projects reinforced key data science skills, including:

**End-to-End ML Pipeline:** Gained practical experience in the full lifecycle of an ML project, from data cleaning and feature engineering to model training and evaluation.

**Tool Proficiency:** Developed a high level of proficiency with professional tools like Python, Google Colab, and GitHub.

**Problem-Solving:** The persistent challenges with data pathing, file structures, and model deployment forced me to develop robust debugging and problem-solving skills.

**Model-Specific Skills:** Gained specialized knowledge in regression modeling with Random Forest and object detection with YOLOv8.

## 8 Future work scope

**Advanced Models:** Explore more advanced models like XGBoost for Project 4 and YOLOv8l for Project 5 to further improve performance.

**Model Optimization:** Investigate techniques to reduce the size and inference time of the YOLOv8 model for deployment on edge devices.

**Interactive Application:** Build a complete interactive web application (like the Gradio demo we discussed) that combines the prediction and detection models into a single, user-friendly tool for farmers.

