Industrial Internship Report on

”Crop Yield Prediction and Weed dectection using Machhine Learning and Deep Learning”

Prepared by

Mucharla Prashanth

Executive Summary

This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner Unconverging 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.

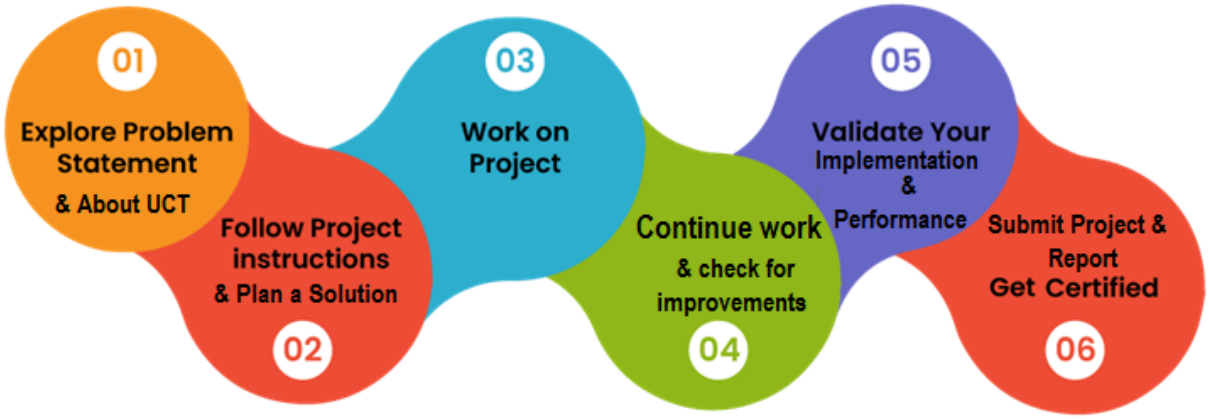
My project was very help full for many users who are really looking for opportunity.

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

Over four weeks, I worked intensively on agricultural data analytics problems, focusing on yield prediction and weed detection. This internship gave me the opportunity to translate classroom learning into solving real-world challenges, enhancing both technical and soft skills. I would like to thank Upskill Campus, The IoT Academy, UCT, and all mentors involved for their continuous guidance.



My Learnings and overall experience.

Thank you Upskill, campus, for providing this opportunity. I have learned many things

while doing the projects I have faced many doubts my LMS is very helped me in completing the project.

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.

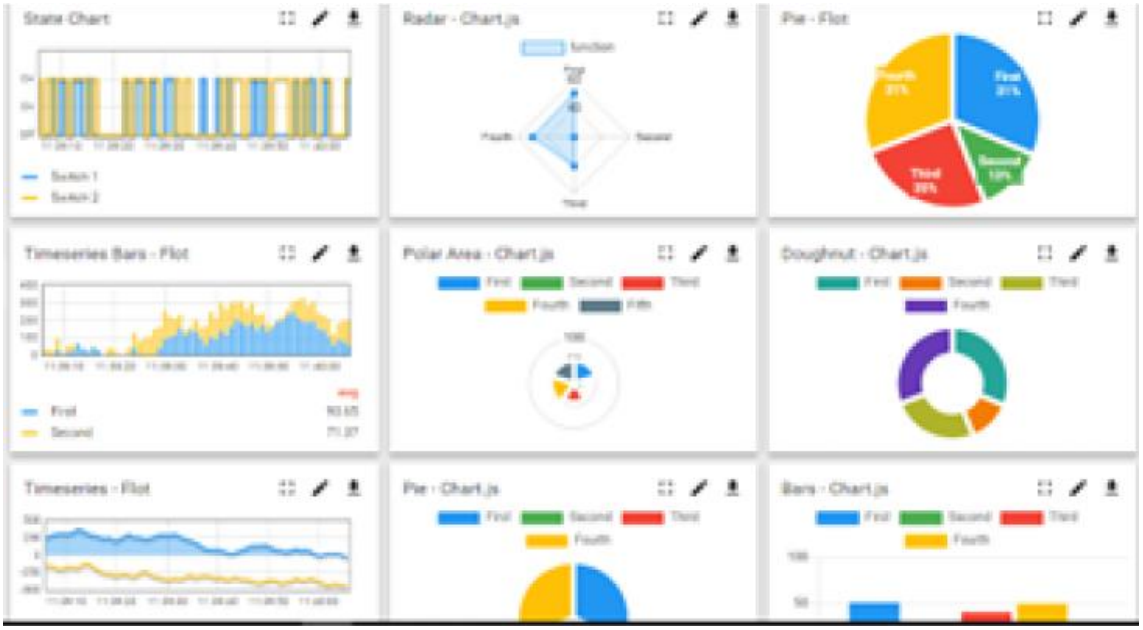
i. UCT IoT Platform (UCT Insight)

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.

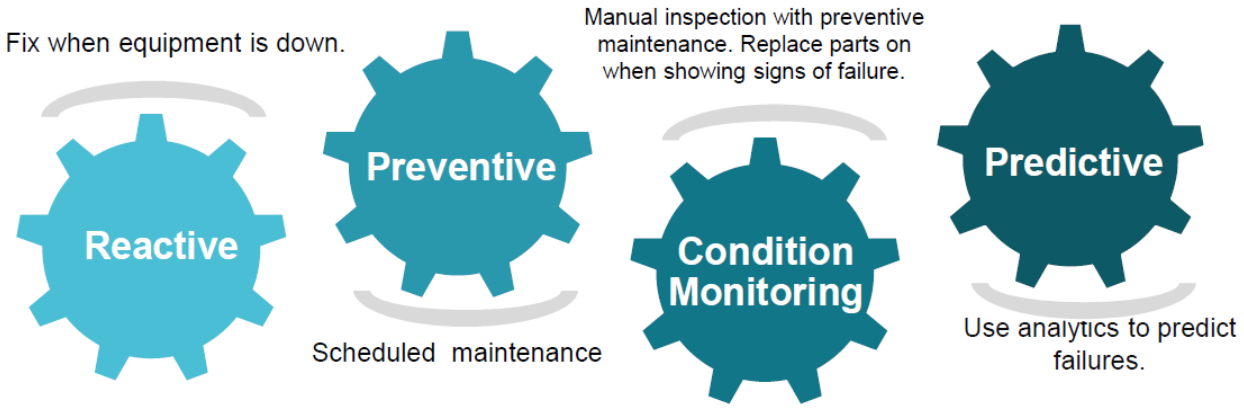
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1	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 technology and providing solution in Agri tech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

iv. Predictive Maintenance

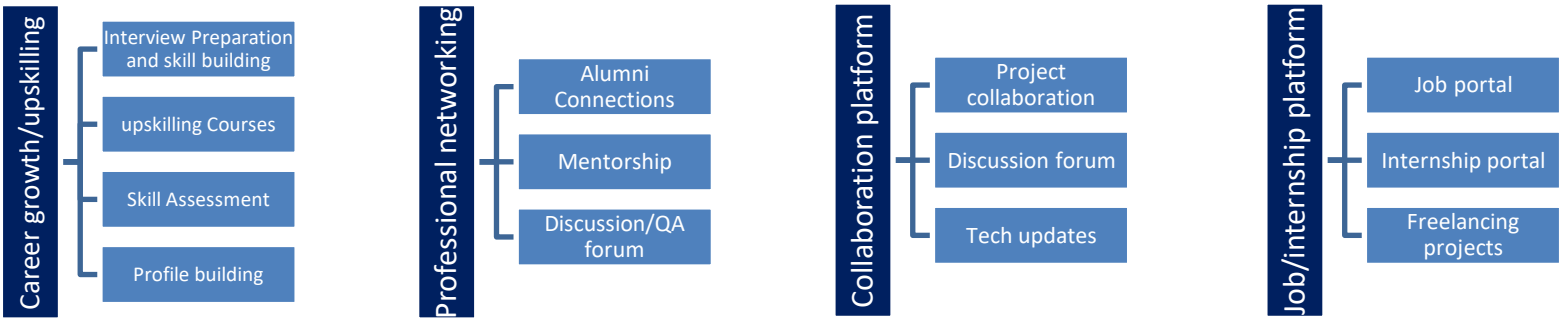
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.



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] Provided agricultural data

[2] Data visualizations from excel and power Bi

[3] Machine learning algorithms from upskill LMS

3 Problem Statement

Farmers encounter significant difficulties in accurately predicting crop yields due to a wide range of dynamic factors, including climate variability, soil characteristics, and cultivation costs. These challenges hinder effective agricultural planning and decision-making. Additionally, the manual detection of weeds is time-consuming and prone to error, often leading to excessive use of herbicides and reduced crop productivity. This project aims to address these issues by developing reliable predictive models using machine learning and deep learning techniques to support both yield forecasting and automated weed detection.

4 Existing and Proposed solution

Existing solutions often lacked scalability and robustness across different agro-climatic conditions. Our proposed solution integrated diverse datasets, applied advanced machine learning models like Random Forest and XGBoost, and leveraged deep learning (CNN, LSTM) for enhanced prediction accuracy. An interactive Streamlit dashboard was developed for end-user accessibility.

4.1 Code submission: crop-yield-prediction-and-weed-detection-using-machine-learning-and-deep-learning/notebooks at main · prashanth-sss/crop-yield-prediction-and-weed-detection-using-machine-learning-and-deep-learning

4.2 Report submission : prashanth-sss/crop-yield-prediction-and-weed-detection-using-machine-learning-and-deep-learning

5 Proposed Design/ Model

Flow:
- Data Collection ➔ Data Preprocessing ➔ Model Training ➔ Model Evaluation ➔ Dashboard Deployment

Models Used:
- Random Forest Regressor
- XGBoost Regressor
- LSTM for time-series data
- CNN for image classification (weed detection)

5.1 High Level Diagram (if applicable)

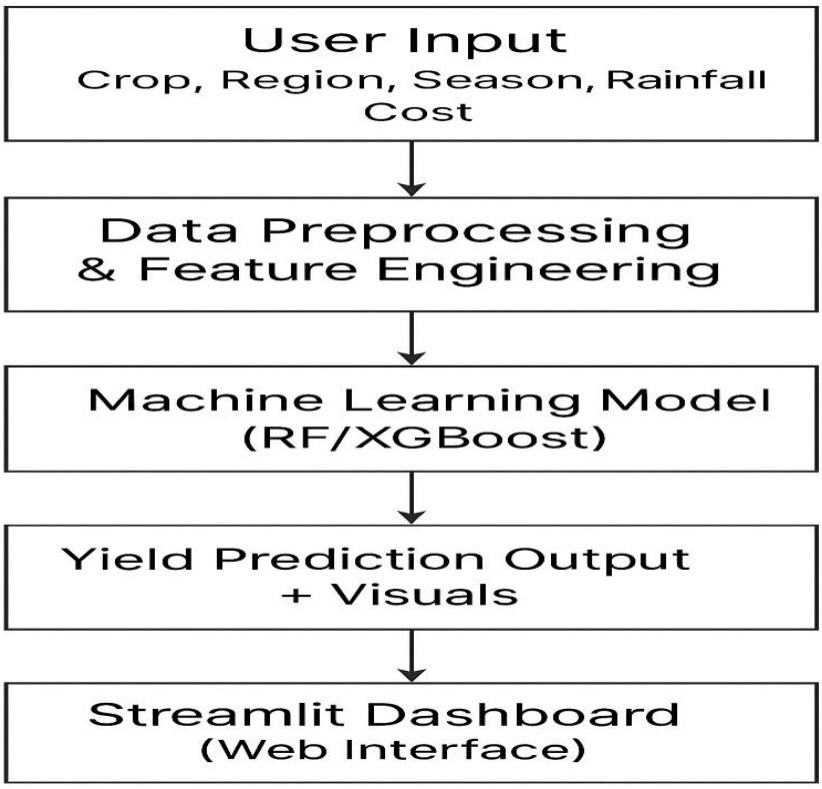
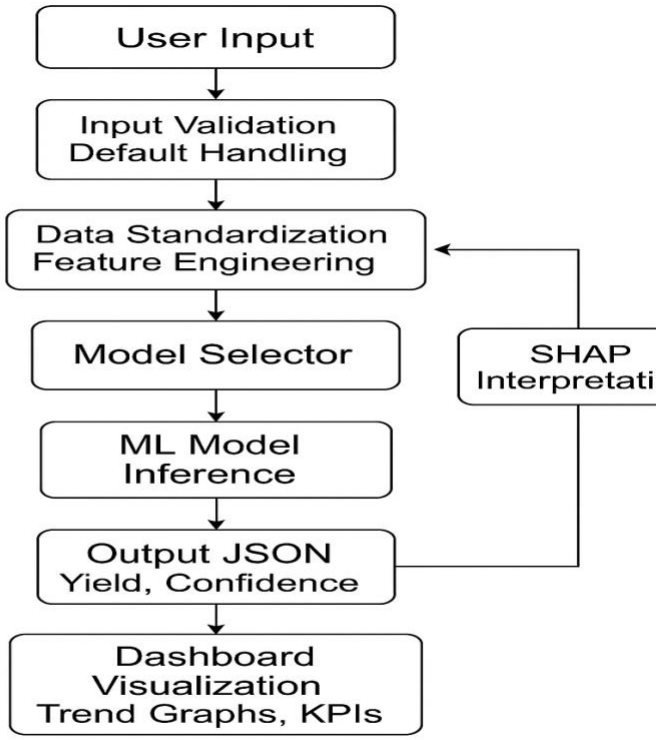


Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

5.2 Low Level Diagram (if applicable)



- **Interfaces in the Project**

1. User Interface (UI) – Streamlit Dashboard

Purpose: To allow users (e.g., farmers, researchers) to input data and view predictions.

Features:

- Input fields for crop type, soil quality, rainfall, region, and cost.
- Interactive charts and trend visualizations (via Plotly).
- Display of predicted yield and model confidence levels.

Data Input Interface

Sources: CSV files, real-world agricultural datasets, potential API integration

Types of Data:

- Tabular data (crop yield history, costs, weather)
- Image data (crop vs. weed classification)
- External data (IoT sensors or satellite data – future scope)

Model Interface (Backend APIs or Function Calls)

Function: Connects the frontend dashboard to the machine learning and deep learning models.

Key Interactions:

- Input preprocessing and formatting
- Model prediction and output interpretation
- Handling model selection dynamically based on input type (ML for tabular, DL for images)

Visualization Interface

Tool: Plotly / Seaborn / Matplotlib

Purpose: To present insights like:

- Crop yield trends
- Feature importance
- Confusion matrix for weed detection

Deployment Interface

- Tools: Streamlit + Local/Cloud Hosting
- Future Add-on: REST API or FastAPI for broader integration or mobile access.

Performance Test

5.3 Test Plan:

- Data split into training and testing sets
- Hyperparameter tuning using Grid Search and Optuna

5.4 Performance Metrics:

- RMSE, MAE, R^2 Score for crop yield prediction
- Accuracy, F1-Score, Precision, recall for weed detection

5.5 Performance Outcome:

- Achieved >85% prediction accuracy.
- Robustness validated across diverse datasets Test Plan/ Test Cases

6 My learnings

This internship significantly enriched my practical knowledge and skills in the field of data science and machine learning. Key takeaways include:

- Gaining hands-on experience in data preprocessing and feature engineering tailored to agricultural datasets.
- Building, training, and evaluating machine learning and deep learning models for real-world applications.
- Deploying models through interactive web applications, enhancing accessibility and user interaction.
- Understanding and overcoming real-world industrial challenges such as data inconsistency, model interpretability, and deployment constraints.

7 Future work scope

While the current system lays a solid foundation, several enhancements can be pursued to expand its real-world impact:

- **Integration of Real-Time Data:** Incorporating satellite imagery and IoT sensor data for more accurate and dynamic yield predictions.
- **Scalability to Diverse Crops and Regions:** Extending the model to cover a wider variety of crops and agro-climatic zones to improve generalizability.
- **Mobile Optimization:** Adapting the solution for mobile platforms to support field-level accessibility for farmers.
- **On-Ground Deployment and Feedback:** Collaborating with agricultural communities to deploy the system in real-world conditions and gather feedback for iterative improvements.