

Industrial Internship Report on
" Crop Yield Prediction "

Prepared by

Yash Modi

Executive Summary

This report provides details of the Industrial Internship facilitated by **Upskill Campus** and **The IoT Academy** in collaboration with **Industrial Partner UniConverge Technologies Pvt Ltd (UCT)**.

This internship focused on a problem statement provided by UCT, and the project was to be completed within a 6-week time frame.

My project was on **Crop Yield Prediction**, where the primary aim was to leverage machine learning to predict crop yields based on various agricultural factors. This project offered a valuable opportunity to address real-world challenges in agriculture and develop a data-driven solution to optimize resource allocation and decision-making.

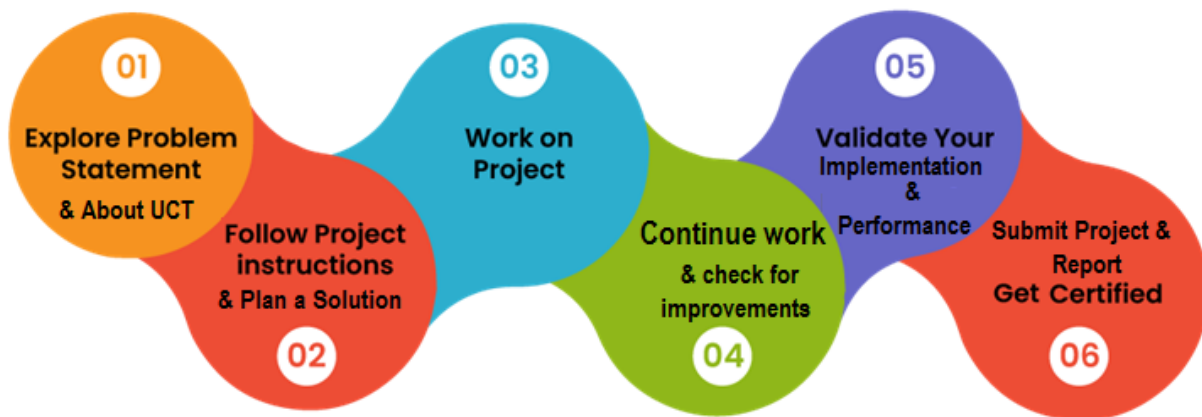
This internship was an excellent platform to gain exposure to industrial problems, design solutions, and implement them using cutting-edge technologies. It was a rewarding experience that bridged the gap between academic knowledge and practical implementation.

TABLE OF CONTENTS

1	Preface	4
2	Introduction	5
2.1	About UniConverge Technologies Pvt Ltd	5
2.2	About upskill Campus	9
2.3	Objective	11
2.4	Reference	Error! Bookmark not defined.
2.5	Glossary	Error! Bookmark not defined.
3	Problem Statement	12
4	Existing and Proposed solution	13
5	Proposed Design/ Model	14
5.1	High Level Diagram (if applicable)	Error! Bookmark not defined.
5.2	Low Level Diagram (if applicable)	Error! Bookmark not defined.
5.3	Interfaces (if applicable)	Error! Bookmark not defined.
6	Performance Test	15
6.1	Test Plan/ Test Cases	15
6.2	Test Procedure	15
6.3	Performance Outcome	15
7	My learnings	16
8	Future work scope	17

1 Preface

This report summarizes the work completed during the 6-week internship on **Crop Yield Prediction**. The internship aimed to provide hands-on experience in solving industrial problems and applying machine learning techniques to real-world datasets.



The problem statement addressed the lack of accurate yield prediction models, and the proposed solution aimed to provide actionable insights for improving agricultural productivity. This internship was well-structured, with each week focusing on specific milestones, including problem understanding, data preprocessing, model development, and final evaluation.

I express my gratitude to **Upskill Campus**, **The IoT Academy**, and **UCT** for this opportunity. Special thanks to my mentors and peers who supported me throughout this journey. To my juniors and peers, I encourage you to take up such internships to enhance your skills and gain real-world exposure.

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.



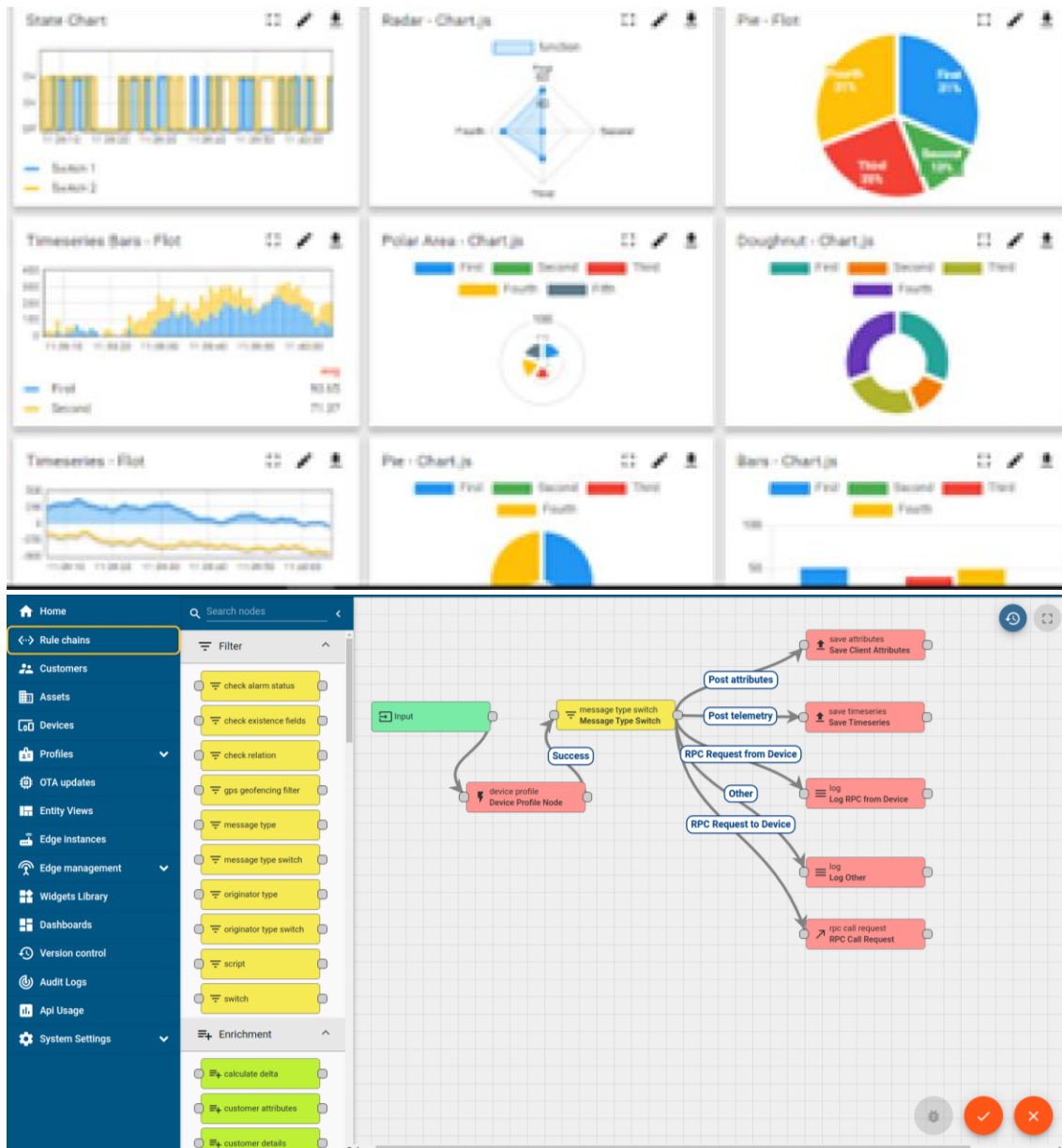
i. UCT IoT Platform ()

UCT Insight 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 unleash 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 want 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.



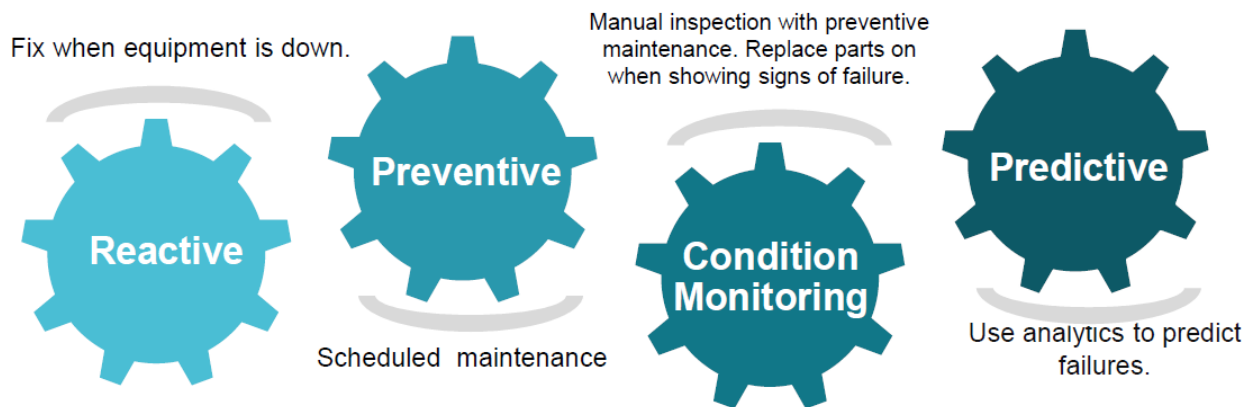


iii. LoRaWAN based Solution

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

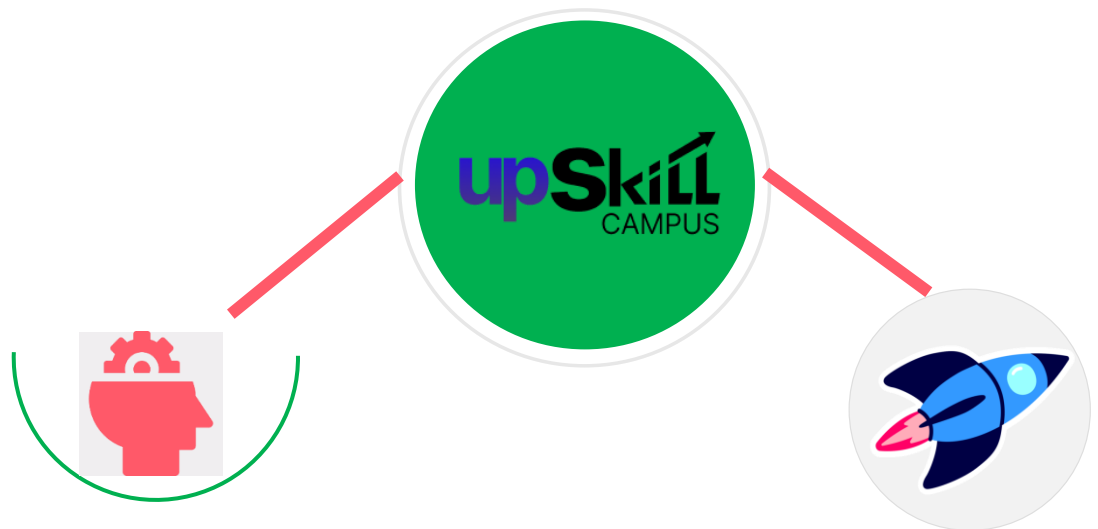
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.

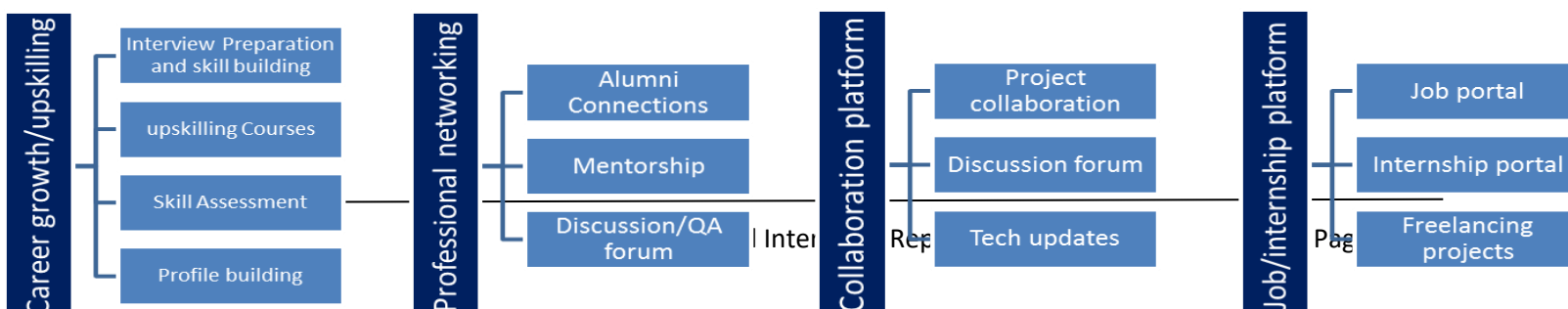
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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.

3 Problem Statement

The assigned problem statement focused on predicting crop yields using machine learning models based on agricultural data. The problem involved developing a model to analyze multiple factors like crop type, state, cost of production, and yield to assist in agricultural planning and decision-making.

4 Existing and Proposed solution

Existing Solutions:

Existing solutions rely on traditional statistical models or expert judgment, which often lack scalability and fail to incorporate complex patterns in the data.

Proposed Solution:

The proposed solution involves using machine learning models, specifically **Random Forest Regressor** and **XGBoost Regressor**, to accurately predict crop yields. The project includes feature engineering, hyperparameter tuning, and model evaluation to ensure reliable performance.

4.1 Code submission (Github link) : https://github.com/YashMmodi-7504/Upskill_Campus/blob/main/CropYieldPrediction.ipynb

4.2 Report submission (Github link) : https://github.com/YashMmodi-7504/Upskill_Campus/blob/main/CropYieldPrediction_Yash_USC_UCT.pdf

5 Proposed Design/ Model

Given more details about design flow of your solution. This is applicable for all domains. DS/ML Students can cover it after they have their algorithm implementation. There is always a start, intermediate stages and then final outcome.

6 Performance Test

The performance test evaluated the effectiveness of the **Crop Yield Prediction** model by applying it to unseen data and measuring its predictive accuracy using metrics like **R²**, **RMSE**, and **MAE**. The test results confirmed the superior performance of the **Random Forest Regressor** ($R^2 = 0.9463$), compared to the **XGBoost Regressor** ($R^2 = 0.8583$). Residual analysis and cross-validation ensured the model's robustness and reliability. These evaluations highlighted Random Forest as the best-suited model for accurate crop yield predictions.

6.1 Test Plan/ Test Cases

- Validate performance on unseen test data.
- Measure metrics such as R^2 , RMSE, and MAE.

6.2 Test Procedure

- Split data into training and testing sets.
- Evaluate models on both datasets and conduct cross-validation.

6.3 Performance Outcome

- **Random Forest R^2 : 0.9463**
- **XGBoost R^2 : 0.8583**

7 My learnings

- **Technical Skills:** Enhanced knowledge of machine learning algorithms, feature engineering, and hyperparameter tuning.
- **Problem-Solving:** Learned to address real-world challenges with scalable solutions.
- **Communication:** Improved ability to document and present findings effectively.

8 Future work scope

- **Deployment:** Deploy the model on cloud platforms for wider accessibility.
- **Additional Features:** Integrate weather and soil data to improve predictions.
- **Scalability:** Expand the model to cover more crops and regions.