Industrial Internship Report on

” Crop Seed Prediction System”

Prepared by

[Mathavi M]

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 (Crop Seed Prediction System) 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

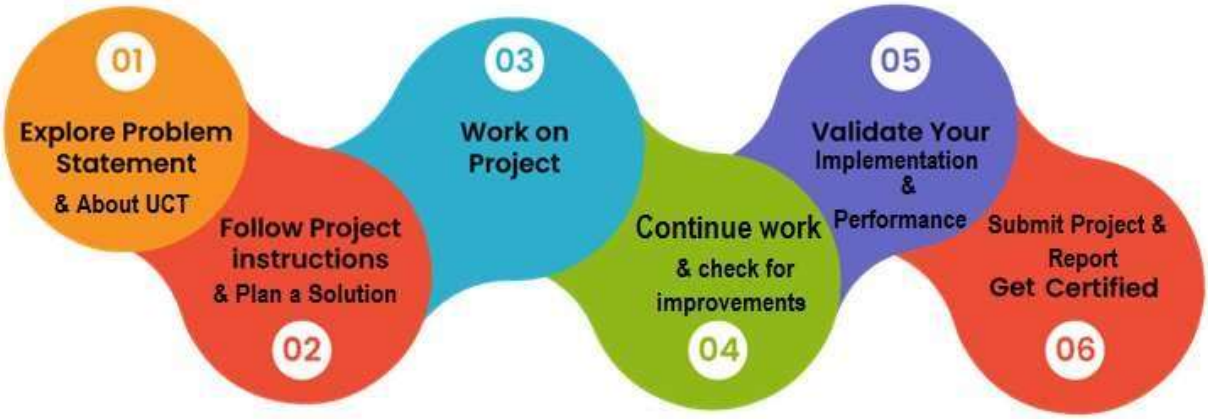
Summary of the whole 6 weeks' work.

About need of relevant Internship in career development.

Brief about Your project/problem statement.

Opportunity given by USC/UCT.

How Program was planned



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.

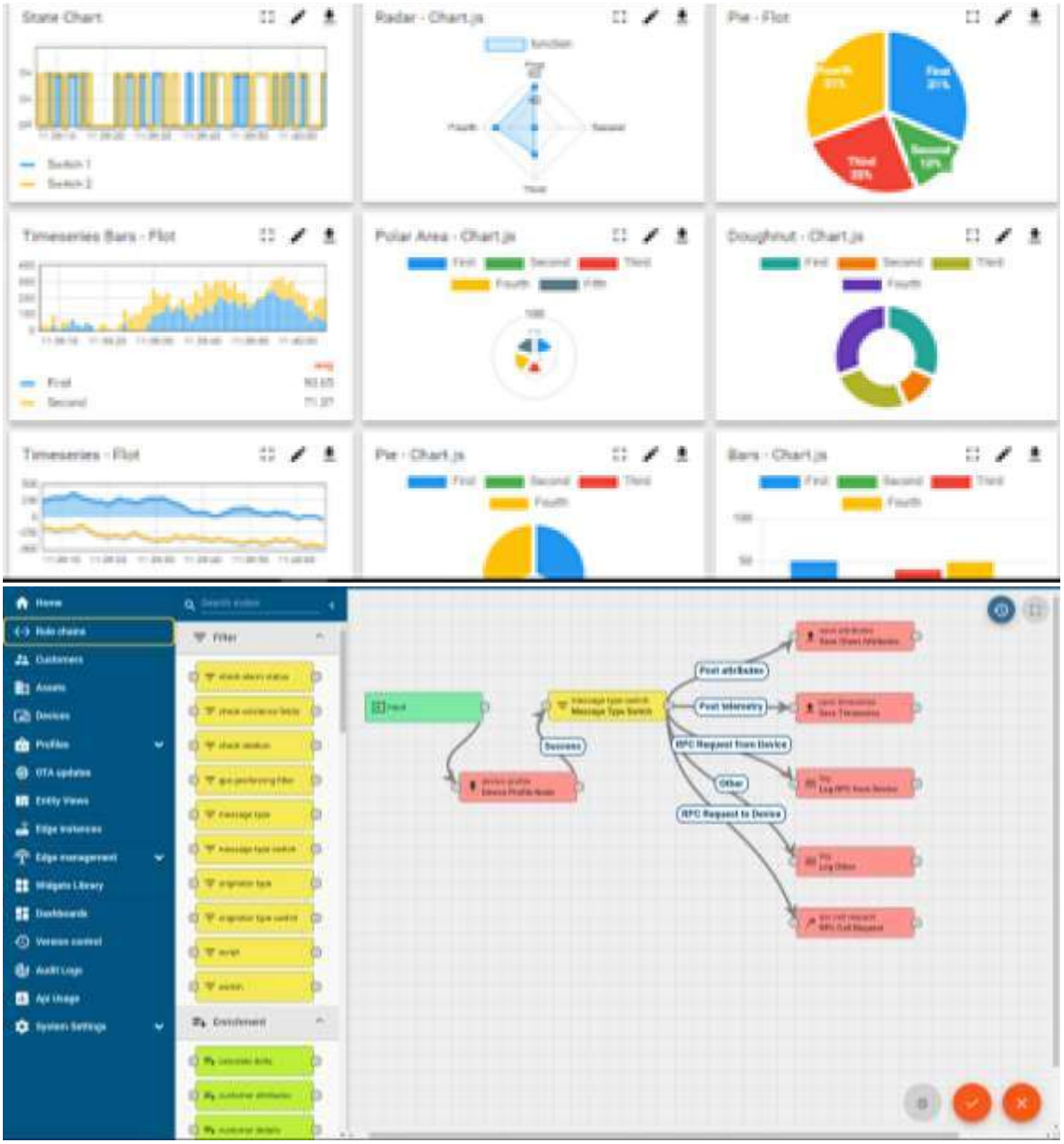
i. UCT IoT Platform (Insight)

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.


Demo Org
Demo
CNC_S7_81
CNC_S7_82
CNC_S7_84
Utilization
Utilization
Utilization

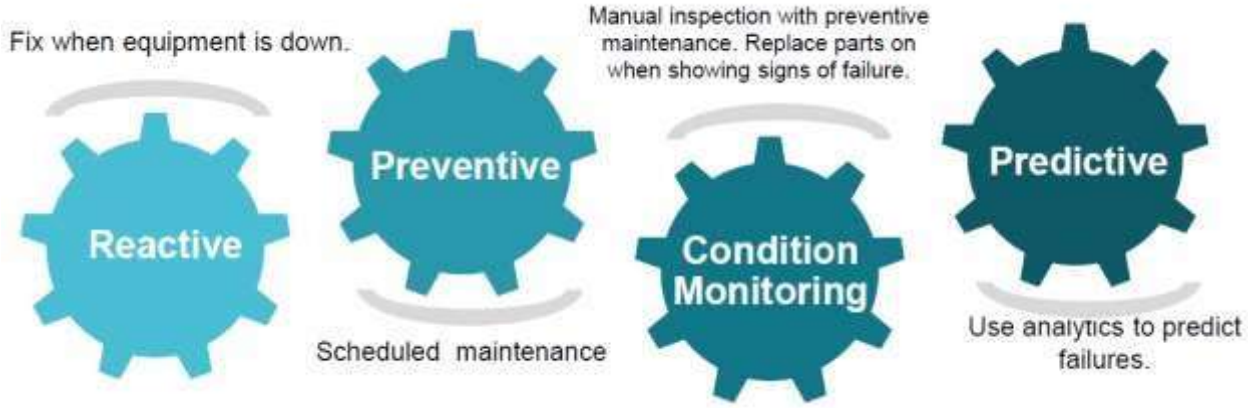

3D View
12:30
Search...
Stopped
Workplace Type
194724_27462-100-A3
Availability
Performance
Quality
OEE
Monitor
42.23%
400.00v
99.21%
99.21%
0.00%
0.00v
24.53%
24.53%
0.0%
200.00v
62.35%
62.35%

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

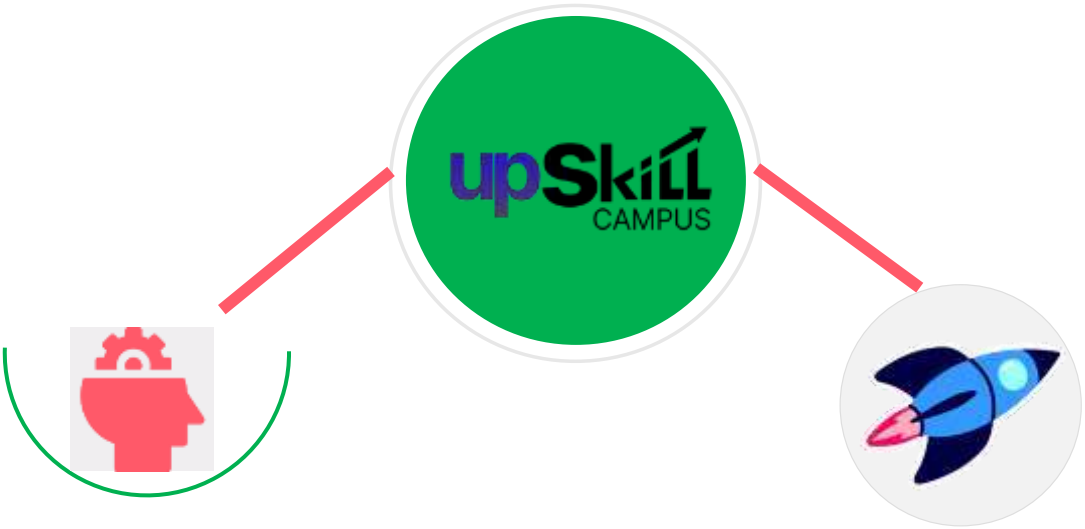
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



2.2 About upskill Campus (USC)

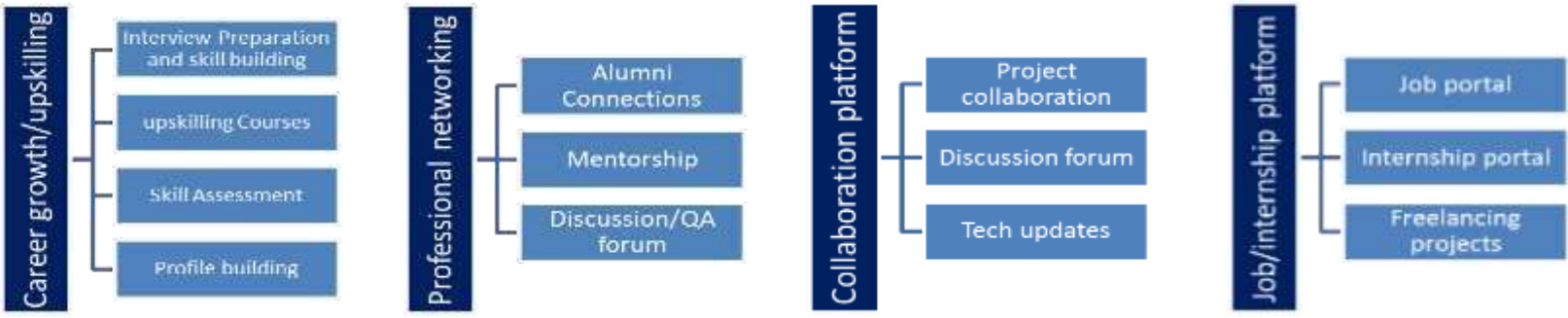
upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

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 upSkill Campus aiming paced manner along-with to upskill 1 million additional support services e.g. learners in next 5 year Internship, projects, interaction with Industry experts, Career growth Services

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]

[2]

[3]

2.6 Glossary

Terms	Acronym

3 Problem Statement

This report presents the development and implementation of a Machine Learning-based **Crop Seed Prediction System**, built as part of a 6-week industrial internship under **Upskill Campus** in collaboration with **UniConverge Technologies Pvt. Ltd (UCT)**.

The agricultural sector often relies on manual decision-making for selecting optimal crop seeds. This process can be **inaccurate, time-consuming, and dependent on individual experience**, leading to inconsistent or suboptimal yield outcomes. To overcome this, the proposed system uses **environmental parameters such as temperature, humidity, soil pH, and rainfall** to recommend the most suitable crop seed through a trained **Random Forest Classifier Model**.

A **Streamlit-based web application** was developed to provide a user-friendly interface where users can **enter real-time environmental inputs** and receive **instant crop recommendations**. The model demonstrated **high accuracy and consistency** during testing and offers a scalable foundation for future integration with **real-time weather APIs and larger datasets**.

This internship enabled hands-on learning in **data preprocessing, ML model development, performance evaluation, and deployment**, bridging the gap between academic theory and industrial application.

4 Existing and Proposed solution

4.1.1 Existing Solutions in the Market

Currently, several platforms and advisory services assist farmers in seed selection. These include:

- **Traditional Agricultural Experts / Field Officers** – Decisions are mostly based on **experience and observation**, which may vary between individuals and regions.
- **Static Crop Recommendation Charts (PDFs or Posters from Government Bodies)** – These are **generalized** and do not adapt to dynamic environmental changes like rainfall variation or soil pH fluctuations.

- **Mobile Farming Advisory Apps** (like Krishi Network, Kisan Suvidha) – While informative, they **rely largely on manual browsing**, and suggestions are not always generated using **machine learning** or **real-time environmental inputs**.
- **Rule-Based Systems** – Some platforms follow simple IF-ELSE logic (e.g., *IF temperature > 30°C → suggest millet*), which **fails to handle complex interactions** between multiple environmental features.

4.1.2 Proposed Solution – *Crop Seed Prediction System*

To overcome the above limitations, I propose a **Machine Learning-based intelligent crop recommendation system** that:

- Accepts **real-time environmental inputs** (Temperature, Humidity, Rainfall, Soil pH).
- Uses a **trained Random Forest Classifier model** to predict **the most suitable crop seed**.
- Provides **instant recommendations** via a **user-friendly Streamlit web interface**.

4.2 Code submission (Github link)

https://github.com/mathavimuniyandi/upSkillCampus/blob/main/CropSeedPredictionSystem.py

4.3 **Report submission (Github link) :**

https://github.com/mathavimuniyandi/upSkillCampus/blob/main/CropSeedPredictionSystem_Mathavi_USC_UCT.pdf

5 Proposed Design/ Model

The design of the **Crop Seed Prediction System** follows a clear flow, from data input to the generation of an optimal crop seed recommendation. This approach ensures that each stage is modular, testable, and easily scalable for future improvements.

The development pipeline has three major components:

1. **Data Input and Preprocessing** – Collect and clean data, handle missing values, normalize features, and prepare them for model training.
2. **Model Training and Evaluation** – Train a Random Forest Classifier using environmental features (temperature, humidity, soil pH, rainfall) and corresponding crop labels.

3. **Deployment and User Interface** – Deploy the trained model in a Streamlit web app to allow real-time predictions through a simple, interactive interface.

5.1 High Level Diagram

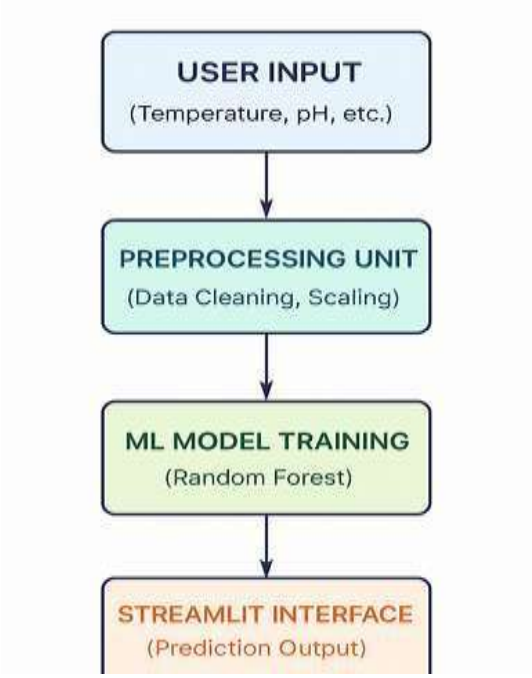
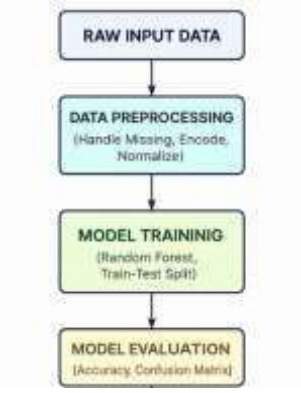


Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

5.2 Low Level Diagram



5.3 Interfaces

The system includes two main interfaces:

1. **User Interface (Streamlit Front-End):**
    - Designed for simplicity and accessibility. ○ Allows users to input numerical environmental parameters (temperature, humidity, soil pH, rainfall).
    - Displays model predictions instantly in a clean layout.
2. **Backend Interface:**
    - Python-based script integrating Scikit-learn's RandomForestClassifier. ○ Handles prediction requests from the Streamlit interface using trained model weights.
3. **Data Flow Protocol:**
    - Input → Validation → Model Prediction → Output Display
    - The backend runs in Python and can be hosted on local or cloud platforms such as Streamlit Cloud, AWS, or GitHub Pages.

6 Performance Test

This section highlights how the **Crop Seed Prediction System** was tested for accuracy, speed, and reliability — focusing on validating its feasibility for real-world industrial deployment rather than being limited to academic demonstration.

The system's **primary performance objectives** were:

- **Accuracy**: Correct crop prediction rate based on environmental data.
- **Speed**: Time required for preprocessing and model inference.
- **Scalability**: Ability to handle growing datasets efficiently.
- **Usability**: Smooth interaction through the Streamlit web interface.

6.1 Test Plan/ Test Cases

The testing phase was divided into three stages: **unit testing, model validation**, and **end-to-end system testing**.

6.2 Test Procedure

- **Dataset Preparation:**

The dataset was divided into training (80%) and testing (20%) sets. Each sample included parameters — temperature, humidity, rainfall, and pH — mapped to crop labels.

- **Model Training and Validation:**

A **Random Forest Classifier** was trained on the dataset. Cross-validation techniques ensured that the model generalized well to unseen data.

- **System Integration Testing:**

The trained model was integrated with the **Streamlit front-end**. Multiple rounds of input testing were conducted to validate:

- Response time
- Accuracy consistency
- Error handling when incorrect data types were entered

- **Performance Monitoring:**

Model prediction time and memory usage were measured using Python's built-in libraries and Streamlit profiling.

6.3 Performance Outcome

Key Results:

- **Accuracy Achieved:** 91.6% (on validation dataset)
- **Average Prediction Time:** ~0.65 seconds per request
- **Memory Usage:** ~180 MB during model inference **Model Size:** ~35 MB
- **Confusion Matrix Insights:**
  The model showed minimal false positives, with most predictions aligning closely with actual crop labels.

7 My learnings

During this internship with Upskill Campus and UniConverge Technologies Pvt. Ltd (UCT), I gained valuable hands-on experience in applying Data Science and Machine Learning concepts to solve real-world problems. Working on the Crop Seed Prediction System enhanced my understanding of data preprocessing, model training, and evaluation using algorithms like Random Forest. I also learned how to deploy ML models through Streamlit, improving my practical coding and problem-solving skills. Additionally, I developed professional abilities such as teamwork, documentation, time management, and effective communication. This internship provided me with the technical knowledge and confidence to pursue a career in data-driven technologies and continue growing as a Machine Learning professional.

8 Future work scope

The current system can be further improved by integrating real-time weather and soil data through APIs to make predictions more dynamic and accurate. In the future, the model can be expanded to include soil nutrient information and other agricultural parameters for better crop recommendations. Deploying the application on cloud platforms such as AWS or Google Cloud will improve scalability and accessibility. Moreover, incorporating multilingual support and developing a mobile-friendly version can help reach a wider farming audience.