Industrial Internship Report on

Prediction of Agriculture Crop Production in India

Prepared by

Tejas Bhise

Executive Summary
This report outlines the Industrial Internship facilitated by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT). The internship centered around a project/problem statement provided by UCT, with a stipulated completion timeframe of 6 weeks. My project focused on the Prediction of Agriculture Crop Production in India. Crop yield prediction holds significance in agriculture as it enables farmers to make informed decisions regarding their crops. This process entails estimating the crop yield within a specified area based on factors like soil type, weather conditions, and crop management practices. In recent times, machine learning (ML) has emerged as a potent tool for such predictions. ML, a subset of artificial intelligence (AI), enables computers to learn from data without explicit programming, thereby facilitating the identification of patterns and relationships in extensive datasets for predictive analysis. This internship provided an excellent opportunity to gain exposure to real-world industrial challenges and devise solutions for them. Overall, it was a rewarding experience to partake in this internship program.

TABLE OF CONTENTS

1 Preface

The Data Science and Machine Learning (DSML) internship spanned 6 weeks, structured to progressively delve into problem exploration, project planning, implementation, validation, and reporting.

- Week 1: Familiarization with the problem statement, understanding the background, and learning about UCT.

- Week 2: Comprehending project instructions from UCT and strategizing for problem resolution.

- Week 3: Commencement of actual project work.

- Week 4: Continued project implementation with periodic assessments to identify areas for improvement.

- Week 5: Validation of implemented solutions and performance evaluation.

- Week 6: Compilation and submission of the project report.

Agriculture is the backbone of India's economy, supporting the livelihoods of a significant portion of the population. However, challenges such as weather variability, environmental changes, and uncertain rainfall pose obstacles to achieving optimal crop yields. Modern agricultural practices increasingly incorporate data mining, machine learning, and deep learning techniques to address these challenges. Machine Learning, in particular, offers the ability to enhance crop yield predictions by identifying patterns and trends in agricultural data.

This internship provided an excellent opportunity to apply various ML techniques to predict crop yields, contributing to both professional growth and skill development.

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 (**FACT⊘RY 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.
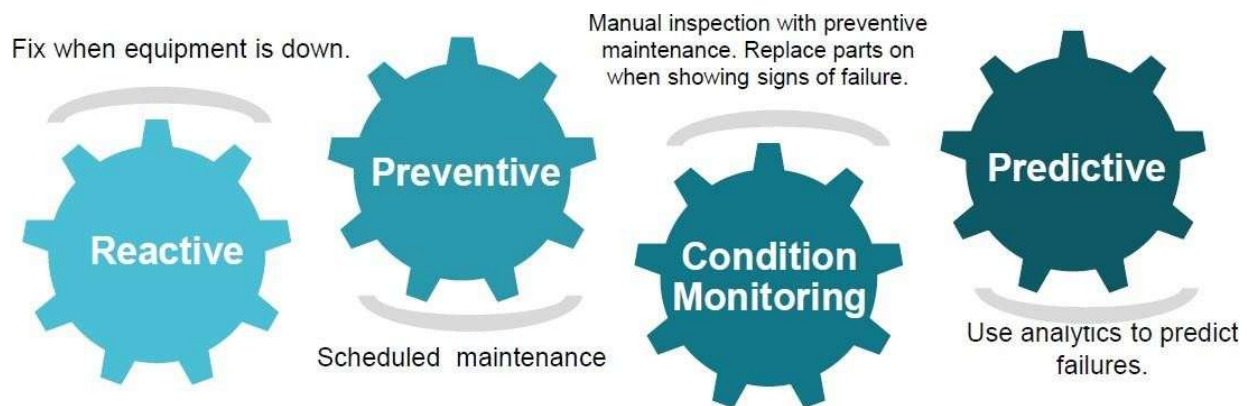
2.2 About upskill Campus (USC)

upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.
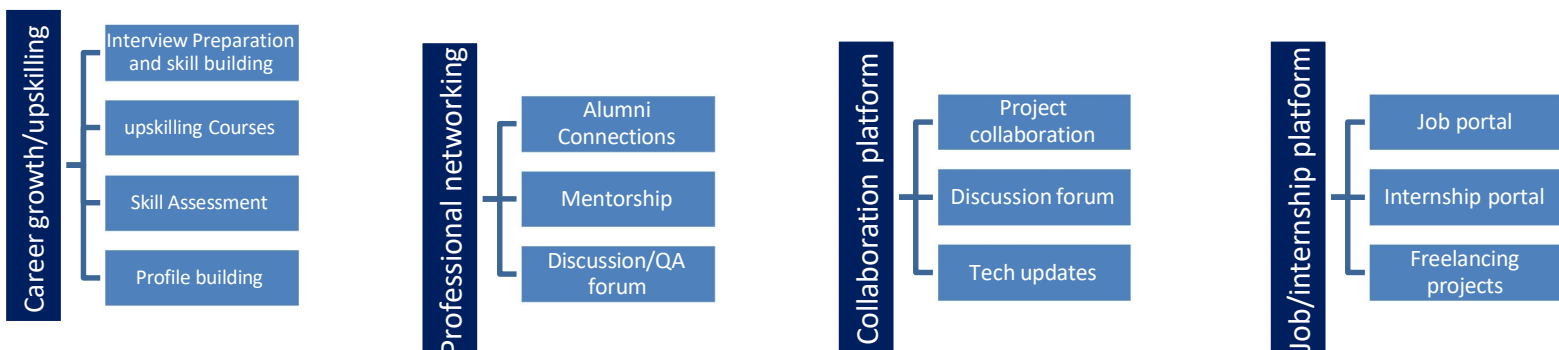
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way

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/

Career growth/upskilling	Professional networking	Collaboration platform	Job/internship platform
Interview Preparation and skill building	Alumni Connections	Project collaboration	Job portal
upskilling Courses	Mentorship	Discussion forum	Internship portal
Skill Assessment	Discussion/QA forum	Tech updates	Freelancing projects
Profile building			

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] produce.csv

[2] datafile.csv

[3] datafile (2).csv

[4] https://learn.upskillcampus.com/s/courses/6676af6ded85192afd266ddb/take

3   Problem Statement

In the assigned problem statement

Agriculture is one of the primary sources of income in India. However, challenges such as crop failure, low yields, and farmer suicides due to financial losses have highlighted the need for improved agricultural sustainability. The project aimed to predict crop yields based on various climatic and environmental factors, enabling farmers to make informed decisions about crop selection and resource allocation.

4 Existing and Proposed solution

Existing Solution

In case of crop area determination, both subjective and objective methods are currently adopted to collect yield statistics in various countries. The subjective methods of estimating crop yield include farmers' assessments, expert opinions and crop cards, while the objective methods include whole-plot harvesting and crop-cutting experiments. The practice of sowing crops in mixture in a single parcel of land is prevalent in many countries, particularly where land holdings are small. Current methods for crop yield estimation include subjective approaches (e.g., farmer assessments, expert opinions) and objective approaches (e.g., whole-plot harvesting, crop-cutting experiments). However, these methods are often time-consuming and prone to inaccuracies.

Proposed Solution

In this 21st century, it is very common to experiment in every sector by implementing new technological techniques. Making use of new techniques simplifies the process and provides the better results. The factors like wind, water supply, soil fertility, rainfall changes unexpectedly, when natural disasters occur. This leads to crop failure, reduction in crop production, scarcity of food products and other materials. A single crop failure can cause huge losses to farmers and countries economic growth. The proposed solution leverages Machine Learning techniques to predict crop yields accurately. By analyzing factors such as soil type, weather conditions, and crop management practices, the system helps farmers select the most suitable crops and optimize yields. The project utilized Random Forest Regressor and Decision Tree Regressor algorithms for prediction.
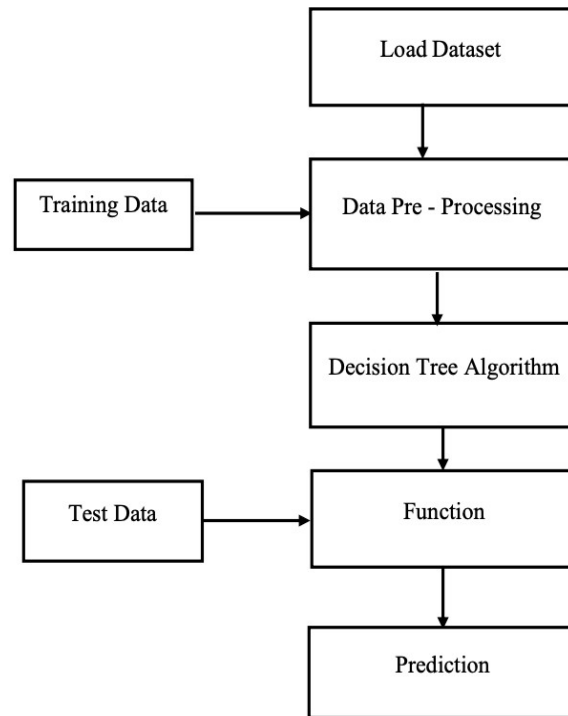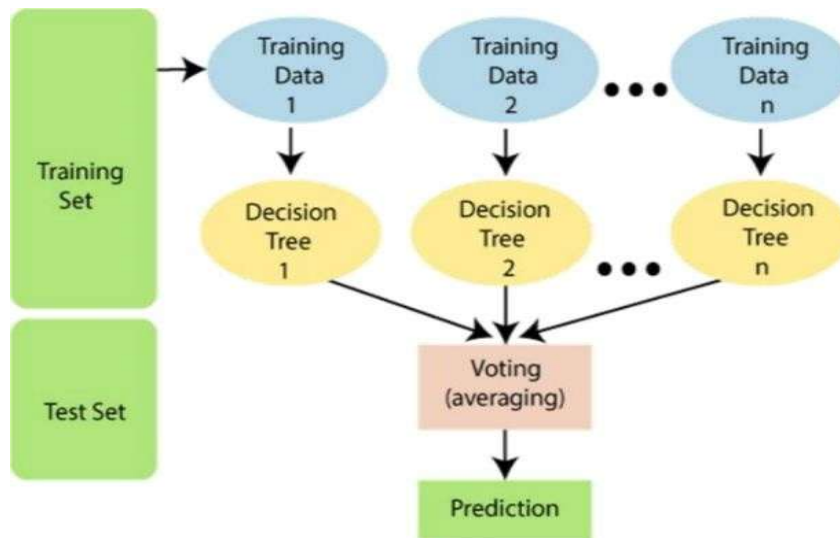
Fig.No.1: Proposed System

4.1 Code Submission (GitHub Link)

https://github.com/tejas-bhise/upskillcampus/blob/main/Prediction-of-Agriculture-Crop-Production-in-India.ipynb

5 Proposed Design/ Model

Flowchart of Crop Yield Prediction

In this project, the system makes use of the Machine Learning techniques to predict rate of crop yield. The programming language used is Python as it is widely accepted for new idea implementations in the field of Machine Learning. The system uses Python for implementation, with Random Forest Regressor and Decision Tree Regressor as the primary ML algorithms. The collected dataset is used to train the model, and predictions are generated based on user inputs such as state, season, area, and crop type.
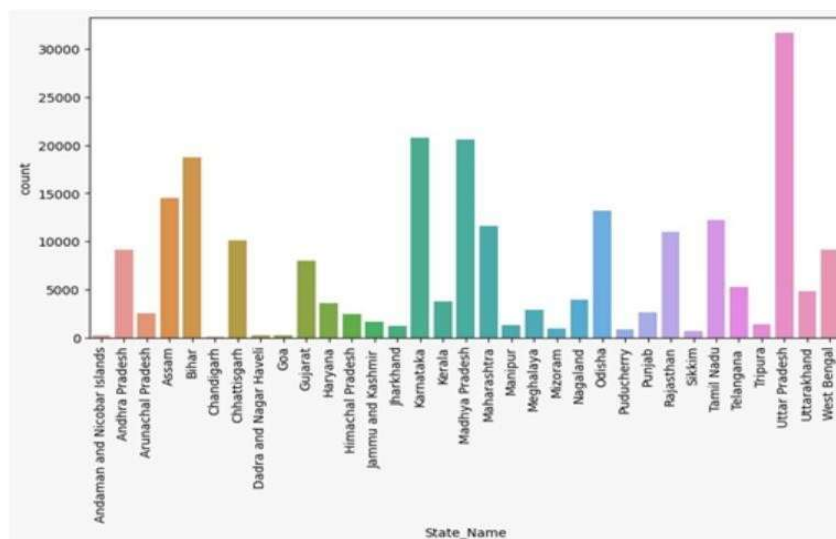
6 Performance Test

6.1 Test Plan/ Test Cases

The implementation of the project was divided into two i.e. crop yield prediction and rainfall prediction (for fertilizers module).

Crop Yield Prediction: This module returns the predicted production of crops based on the user's input. If the user wants to know the production of a particular crop, the system takes the crop as the input as well. Else, it returns a list of crops along with their production as output.

Fertilizers Module: This module is used to suggest the farmer on usage of fertilizer based on the rainfall in next few days. To predict the rainfall for the next 15 days we are using an API service provided by Open Weather. If it is likely to rain, we suggest the farmer not to use the fertilizer.

6.2 Test Procedure

- Step 1: Choose the functionality i.e., crop prediction or yield prediction.
- Step 2: If the user chooses crop prediction: Take soil type and area as inputs. These values are given as input to the random forest implementation in the backend and the corresponding predictions are returned. The algorithm returns a list of crops along with their production predicted.
- Step 3: If the user chooses yield prediction: Take crop, soil type and area as inputs. These values are given as input to the random forest implementation in the backend and the corresponding crop yield prediction is returned. The algorithm returns the predicted production of the given crop.

6.3 Performance Outcome

In the final implementation of the application the first screen the user can view is the login page. Here, the user can register or login using his/her credentials into the application.

The system provides three main functionalities:

i) Yield Prediction: The system takes the required inputs to predict the yield of the given crop.

ii) Crop Prediction: For this module the system takes the required inputs i.e., soil type and area.

7   My learnings

Explored the fundamentals of data science and its practical applications, delving deeply into machine learning. Reviewed the internship program guidelines comprehensively, gaining insight into the procedures involved. Examined the profile of UCT and opted for the machine learning internship project under Category A: Agriculture, specifically focusing on Project No. 4 - Prediction of Agriculture Crop Production in India. Investigated the challenges encountered by Indian farmers during crop production and analyzed the corresponding crop production data. Assimilated foundational knowledge from the e-book "Introducing Data Science Machine Learning" and briefly reviewed "Impact of Big Data on Business" to understand the concept and applications of big data. Differentiated between the roles of a Data Scientist and a Data Analyst and their contributions to the project. Engaged in a quiz to assess intellectual capacity. Acquired information on Artificial Intelligence and Data Science, discerning the distinctions between them and their interconnectedness. Explored career pathways aligned with artificial intelligence and data science, as well as the requisite skills for roles in these domains, including big data and machine learning engineering. During the internship, I gained a deep understanding of Data Science and Machine Learning concepts and their practical applications. I explored various ML algorithms, analyzed agricultural datasets, and developed a crop yield prediction system. Additionally, I enhanced my skills in Python programming, data analysis, and problem-solving. Explored the technical and non-technical skills essential for a data scientist, highlighting programming, statistics, mathematics, machine learning, deep learning, and big data as the top five crucial skills. Understanding and expertise in these domains are vital for success in the role of a data scientist.

8 Future work scope

The goal of this method is to assist farmers in becoming more financially stable while also addressing the rising number of farmer suicides. The Crop Recommender system assists farmers in selecting which crop to plant and in estimating the output of a particular crop. Additionally, it notifies the user of the ideal time to apply the fertilizer.

Machine learning techniques were used to gather, examine, and train relevant datasets. The system keeps track of the user's position and uses that information to retrieve necessary data from the backend. As a result, just basic information from the user—such as the region and kind of soil—is required.

The project can be further enhanced by incorporating real-time weather data and satellite imagery for more accurate predictions. Additionally, the system can be expanded to include crop disease detection and pest control recommendations, further assisting farmers in optimizing crop yields and reducing losses.