

Industrial Internship Report on Prediction of Agriculture Crop Production in India

**Prepared by
Ajay Raj**

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	3
2	Introduction	4
2.1	About UniConverge Technologies Pvt Ltd.....	4
2.2	Smart factory platform.....	6
2.3	Predictive Maintainance.....	7
2.4	About upskill Campus	8
2.5	Objective	10
2.6	Reference	10
3	Problem Statement	10
4	Existing and Proposed solution	11
5	Code Submission.....	12
6	Proposed Design/ Model	13
7	Performance Test	14
7.1	Test Plan/ Test Cases.....	14
7.2	Test Procedure	14
7.3	Performance Outcome	15
8	My learnings	15
9	Future work scope	16

1 Preface

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

During the 1st week, the focus was on familiarizing oneself with the provided problem statements and understanding their background, alongside learning about UCT. Week 2 involved comprehending and adhering to the project instructions from UCT, while also strategizing for problem resolution.

In the 3rd week, actual project work commenced, followed by continued efforts in Week 4, with periodic assessments to identify areas for improvement. Week 5 was dedicated to validating the implemented solutions and evaluating performance. The final week centered on compiling and submitting the project report for certification.

Agriculture forms 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. To address these challenges, modern agricultural practices incorporate data mining, machine learning, and deep learning techniques. Machine Learning, in particular, offers the ability to enhance crop yield predictions by discerning patterns and trends in agricultural data.

In this internship, various machine learning approaches including Linear Regression, Gradient Boosting Regressor, Random Forest Regressor, Decision Tree Regressor, Polynomial Regression, and Ridge Regression were applied to predict crop yields across different states and crop types.

Participating in this internship provided valuable exposure to industrial challenges and the opportunity to devise and implement solutions. Overall, it was an enriching experience that contributed to 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.

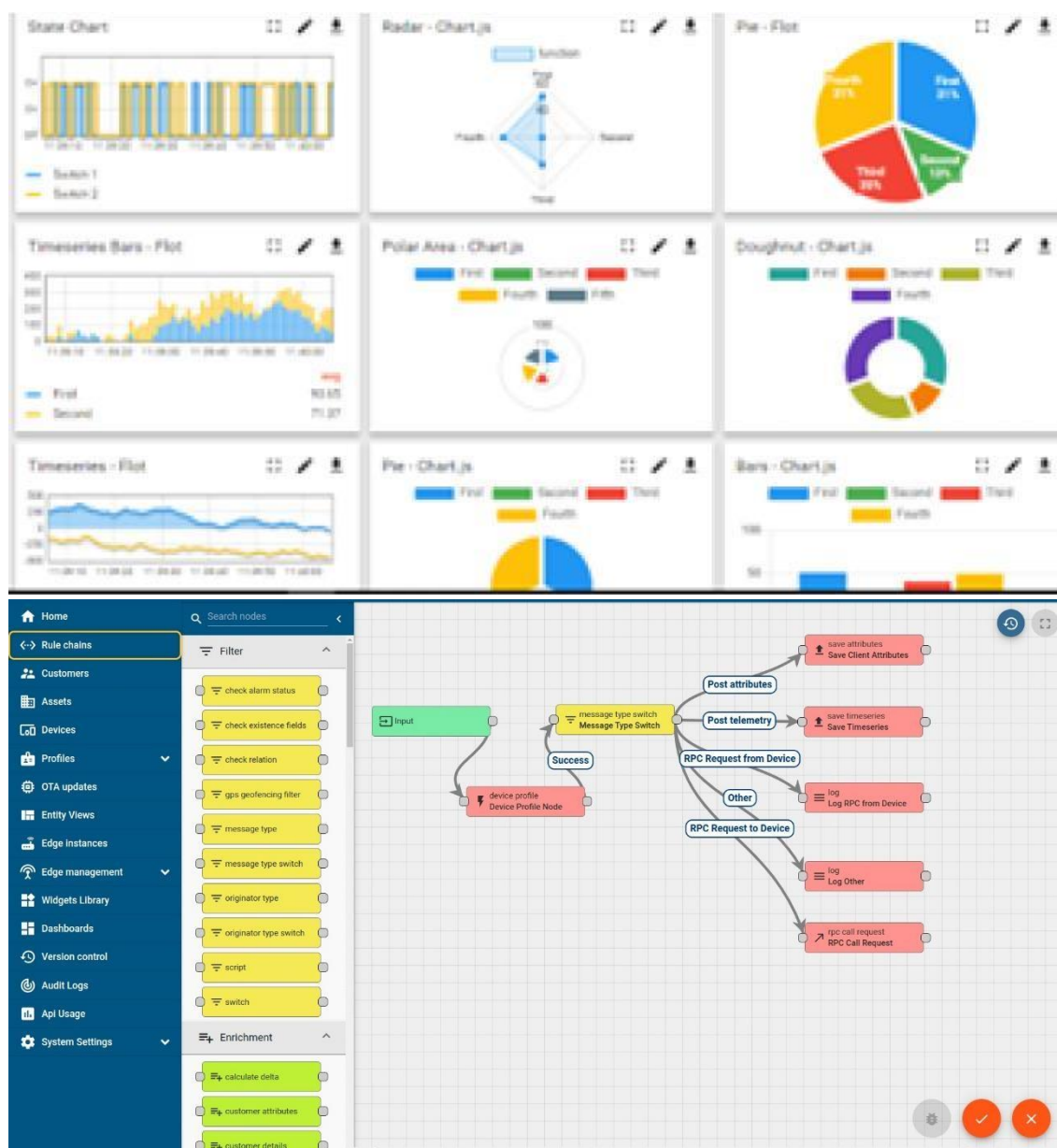


i. UCT IoT Platform ()

UCT Insight 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



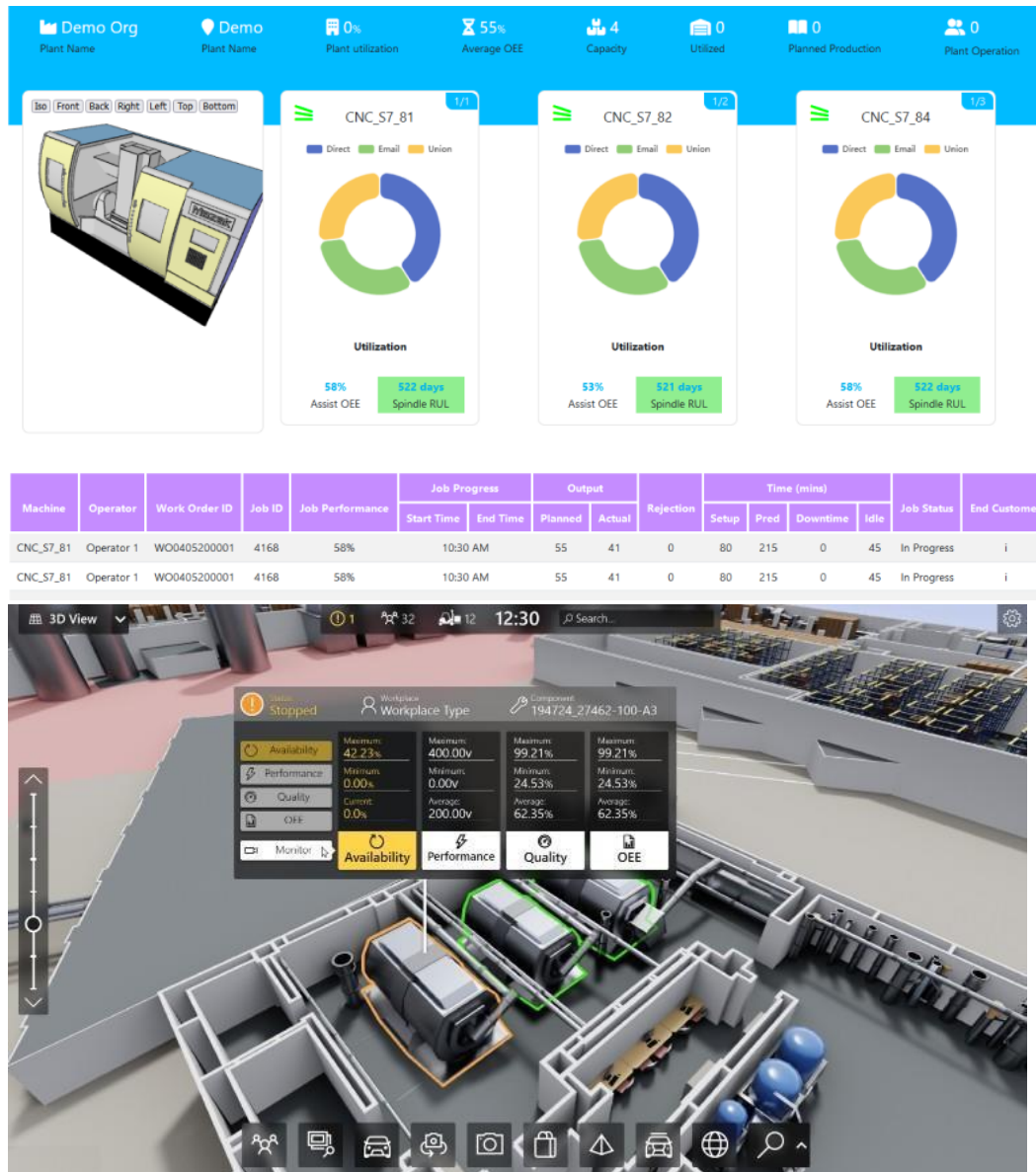
FACTORY WATCH

ii. Smart Factory Platform ()

Factory Watch is a platform designed to meet the needs of smart factories, providing:

- A scalable solution for monitoring production and managing assets.
- Options for Overall Equipment Effectiveness (OEE) and predictive maintenance, with the ability to create digital twins of assets.
- Tools to fully utilize machine-generated data, helping identify and improve key performance indicators (KPIs).
- A modular architecture, allowing users to start with essential services and expand to more advanced solutions as needed.

Factory Watch's unique SaaS model enables users to save time, cut costs, and boost efficiency.

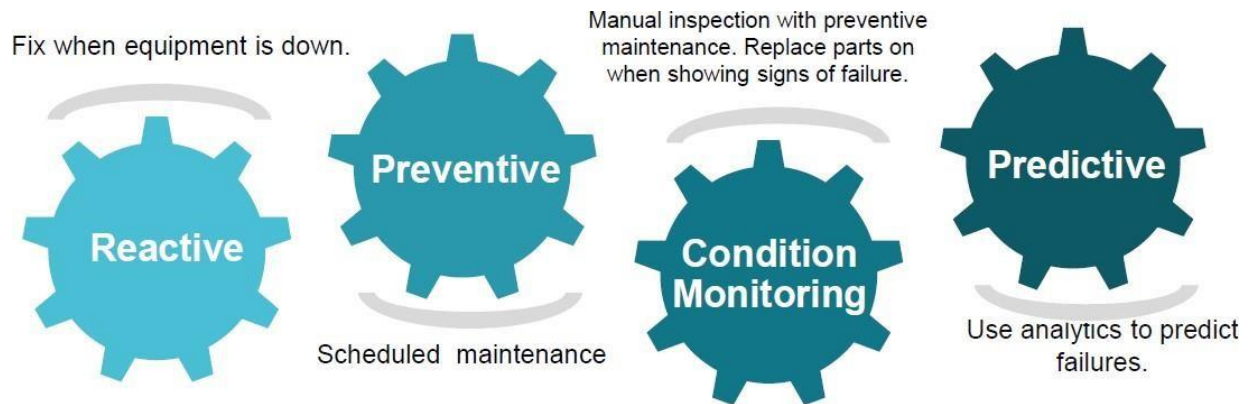


iii. based Solution

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, in collaboration with The IoT Academy and in partnership with Uniconverge Technologies, has ensured the seamless execution of the entire internship process. USC is a career development platform that offers personalized executive coaching in a more affordable, scalable, and measurable manner.

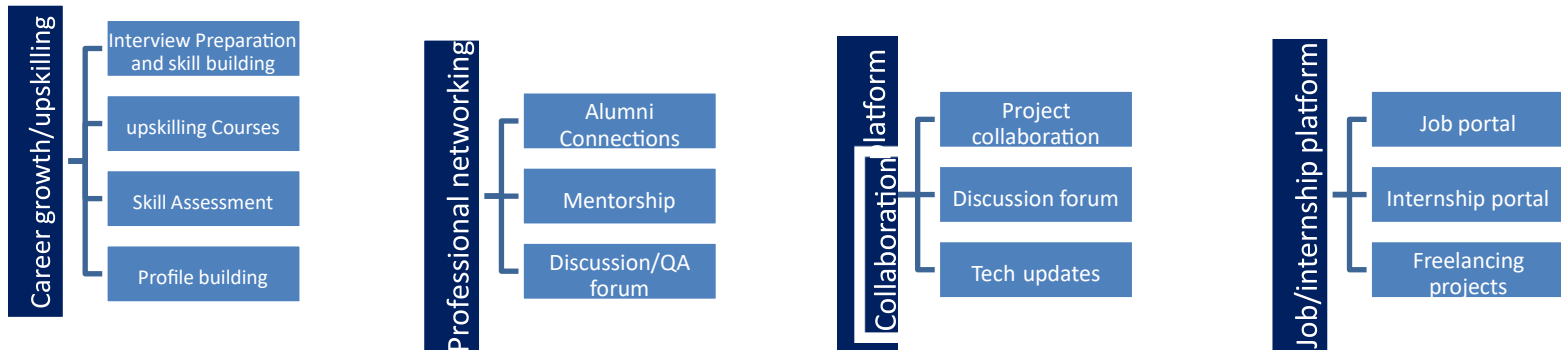


Seeing need of upskilling in self

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 objectives of this internship program were to:

- Obtain practical experience within an industry environment.
- Tackle real-world problems.
- Boost career prospects.
- Expand understanding of our field and its practical uses.
- Promote personal growth, such as better communication and enhanced problem-solving abilities.

2.5 Reference

[1] produce.csv

[2] datafile.csv

[3] datafile (2).csv

[4] <https://learn.upskillcampus.com/s/courses/649195fbe4b0d2807d71596f/take>

3 Problem Statement

In the assigned problem statement

Agriculture is one of the main sources of income in India. There is need to improve the sustainability of agriculture with the rate of increase in suicides of farmer due to crop failure and less yield and losses. Hence, it is a significant contribution towards the economic and agricultural welfare of the countries across the world. The Problem Statement revolves around prediction of yield of crops considering different climatic conditions of India including various attributes. Goal of this project is to help the farmers to choose the suitable crop to grow in order to get the required yield and the profit. Need for the crop yield prediction is very much essential at this point of time for selecting the right crop.

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. The growing of crops in mixtures is a common practice because it protects farmers from adverse 10 weather conditions such as drought, flood, and pest and disease infestation. Further, it enables maximal utilization of the space, moisture and nutrients available in the field. Cultivators usually mix crops that cannot withstand a particular type of weather with another set of crops that thrive under those same conditions.

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. So, there is a desperate need for a new system which can predict the rate of production of crop yield accurately. In order to eradicate all such problems, we have proposed this new system, in which high yielding crop will be selected by considering most influencing parameters. This system helps the farmers to meet their crop yield production. The chances for failure of crops will be very less. In this proposed system, Machine Learning techniques like Random Forest Regressor and Decision Tree Regressor are made used to predict the rate of production of crop yield considering the input parameters like State Name, Season, Area, Crop.

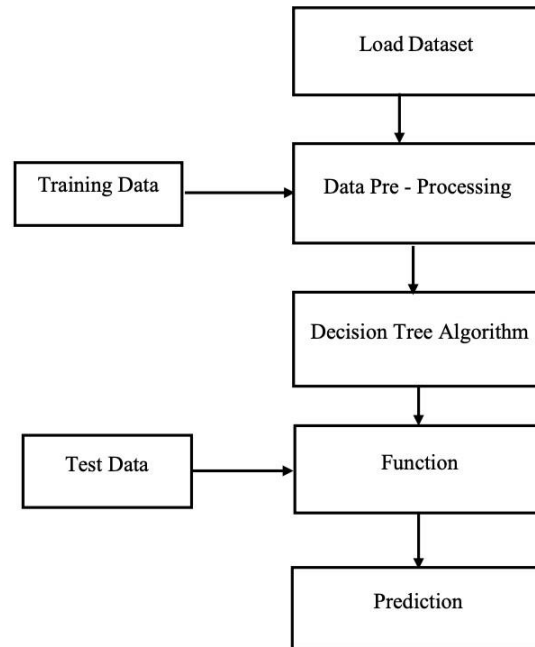
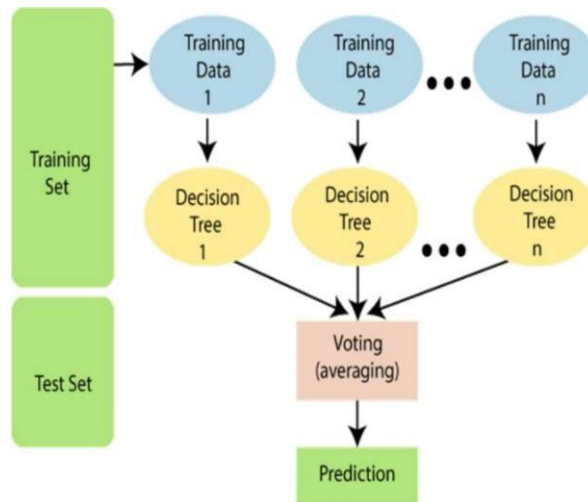


Fig.No.1: Proposed System

4.1 Code Submission (Github Link)

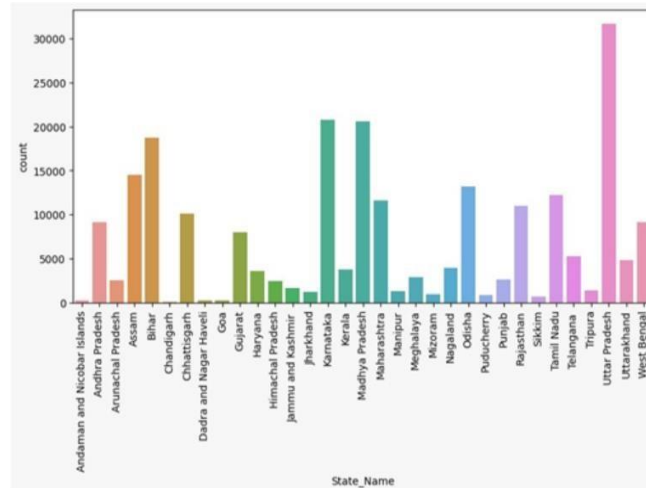
<https://github.com/ajayraj1122/Upskill-Campus-UCT-Prediction-of-Agriculture-Crop-Production-in-India>

5Proposed 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. In this project, collected data set will be uploaded and prediction for crop yield will be generated by applying Machine Learning techniques like Random Forest Regressor and Decision Tree Regressor. The results depend on the information present in the collected data set. Accurate the information about the parameters in the collected datasets, better the results will be.



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 completed application, the first screen users encounter is the login page, where they can either register or log in with their credentials.

The application includes three main functionalities:

1. **Yield Prediction:** Users can enter essential data to predict the yield of a particular crop.
2. **Crop Prediction:** This feature enables users to input details like soil type and area to identify the best crops for their land.

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. Recognized the importance of proficiency in both artificial intelligence and data science, with a realization that specialization in one necessitates a foundational understanding of the other. Acknowledged that, for embarking on machine learning and artificial intelligence, a strong grasp of data analysis is paramount. Briefly

revisited the basics of probability and statistics, covering topics such as sample spaces, random variables, probability distributions, and parametric point estimation. 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

This method aims to enhance farmers' financial stability and mitigate the growing problem of farmer suicides. The Crop Recommender system assists farmers in choosing appropriate crops to cultivate and predicting the yield of these crops. It also informs users about the best times to apply fertilizers.

Machine learning techniques were utilized to gather, analyze, and train on pertinent datasets. The system monitors the user's location and leverages this information to access the required data from the backend. Thus, the user only needs to input basic information, like the region and soil type.