

## Industrial Internship Report on

### **Prediction of Agriculture Crop Production in India**

**Prepared  
by Shiva Posi  
Reddy**

#### *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	<a href="#">Preface</a>	3
2	<a href="#">Introduction</a>	4
1.	<a href="#">About UniConverge Technologies Pvt Ltd</a>	4
2.	About upskill Campus	8
3.	Objective	10
4.	Reference	10
5.	Glossary	
3	<a href="#">Problem Statement</a>	
	<a href="#">Existing and Proposed solution</a>	11 4 12 5
	<a href="#">Proposed Design/ Model</a>	14
6	<a href="#">Performance Test</a>	15
1.	<a href="#">Test Plan/ Test Cases</a>	15



.....15

3. [Performance](#)

[Outcome.....15](#)

7 [My learnings.....16](#)

8 [Future work scope.....17](#)

## 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

### 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

- 
- The screenshot displays the ThingsBoard interface, which is divided into two main sections: a dashboard on the left and a rule chain editor on the right.
- Dashboard Section:**
- State Chart:** A bar chart showing the state of two switches (Switch 1 and Switch 2) over time. The y-axis represents the state (0 or 1), and the x-axis shows time from 11:28:15 to 11:28:30.
  - Radial Chart:** A radial chart showing the state of four switches (Switch 1, Switch 2, Switch 3, and Switch 4) over time. The y-axis represents the state (0 or 1), and the x-axis shows time from 11:28:15 to 11:28:30.
  - Pie Plot:** A pie chart showing the distribution of four categories: First (25%), Second (25%), Third (25%), and Fourth (25%).
  - Timeseries Bars - Plot:** A bar chart showing the state of two switches (Switch 1 and Switch 2) over time. The y-axis represents the state (0 or 1), and the x-axis shows time from 11:28:15 to 11:28:30.
  - Polar Area Chart:** A polar area chart showing the state of four switches (Switch 1, Switch 2, Switch 3, and Switch 4) over time. The y-axis represents the state (0 or 1), and the x-axis shows time from 11:28:15 to 11:28:30.
  - Doughnut Chart:** A doughnut chart showing the distribution of four categories: First (25%), Second (25%), Third (25%), and Fourth (25%).
  - Timeseries - Plot:** A line chart showing the state of two switches (Switch 1 and Switch 2) over time. The y-axis represents the state (0 or 1), and the x-axis shows time from 11:28:15 to 11:28:30.
  - Pie Chart:** A pie chart showing the distribution of four categories: First (25%), Second (25%), Third (25%), and Fourth (25%).
  - Bar Chart:** A bar chart showing the state of two switches (Switch 1 and Switch 2) over time. The y-axis represents the state (0 or 1), and the x-axis shows time from 11:28:15 to 11:28:30.
- Rule Chain Editor Section:**
- Home:** A sidebar menu with options like Home, Rule chains, Customers, Alarms, Devices, Profiles, OTA updates, Entity Views, Edge instances, Edge management, Widgets & Library, Dashboards, Version control, Audit Logs, API Usage, and System Settings.
  - Search rules:** A search bar for finding specific rule chains.
  - Filter:** A dropdown menu for filtering rule chains.
  - Rule Chain:** A visual representation of a rule chain, showing the flow of data and actions. The rule chain starts with a "Post" node, followed by a "device profile" node, then a "Message type switch" node. The "Message type switch" node has four outgoing flows: "Post attributes", "Post telemetry", "RPC Request from Device", and "Other". Each flow leads to a specific action node: "Post attributes" leads to "Post attributes", "Post telemetry" leads to "Post telemetry", "RPC Request from Device" leads to "RPC Request from Device", and "Other" leads to "Log RPC from Device".

## **FACTORY** **WATCH**

### ii. SmartFactory Platform ( )

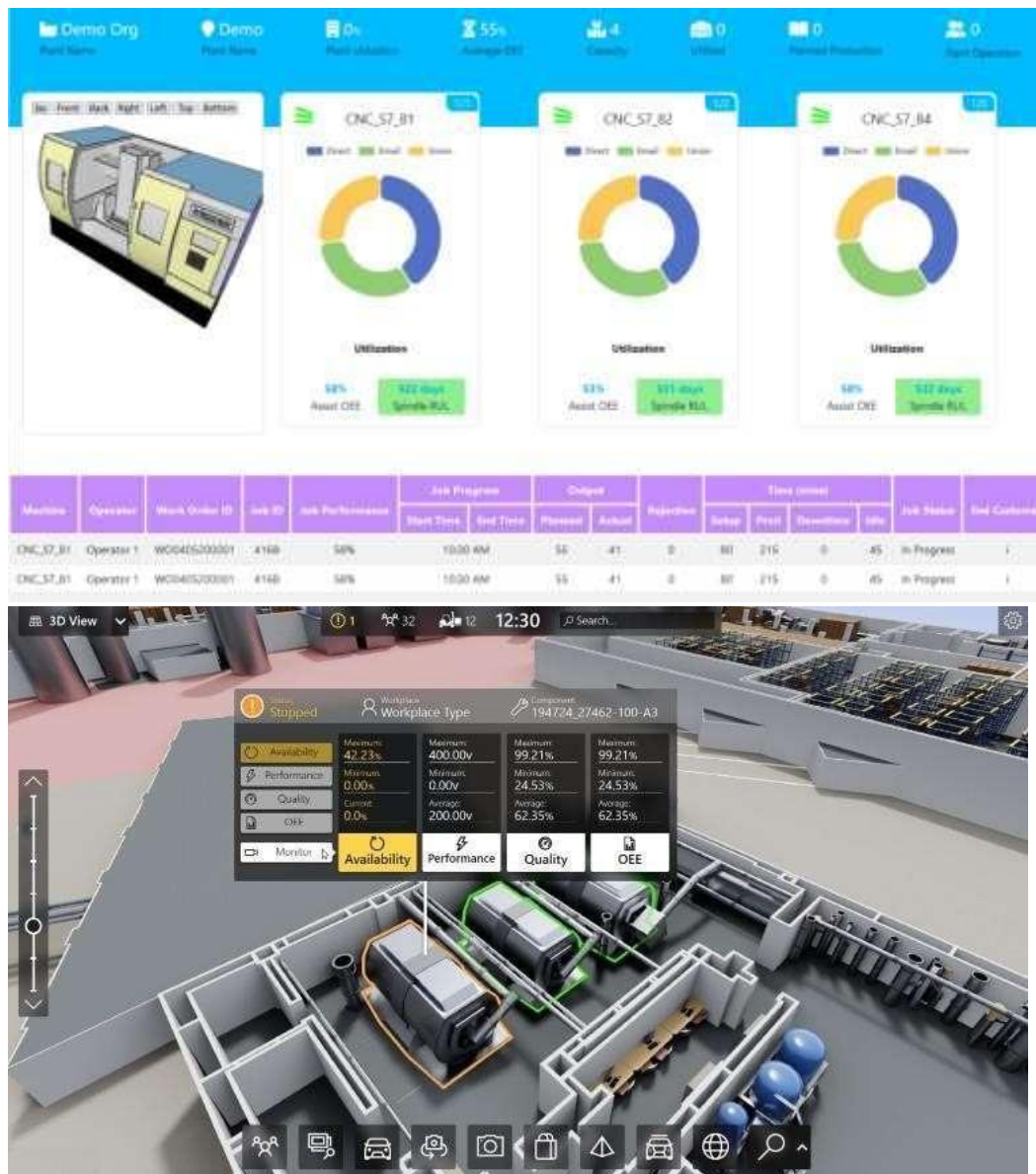
Factory watch is a platform for smart factory needs.

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.





**LoRaWAN™**

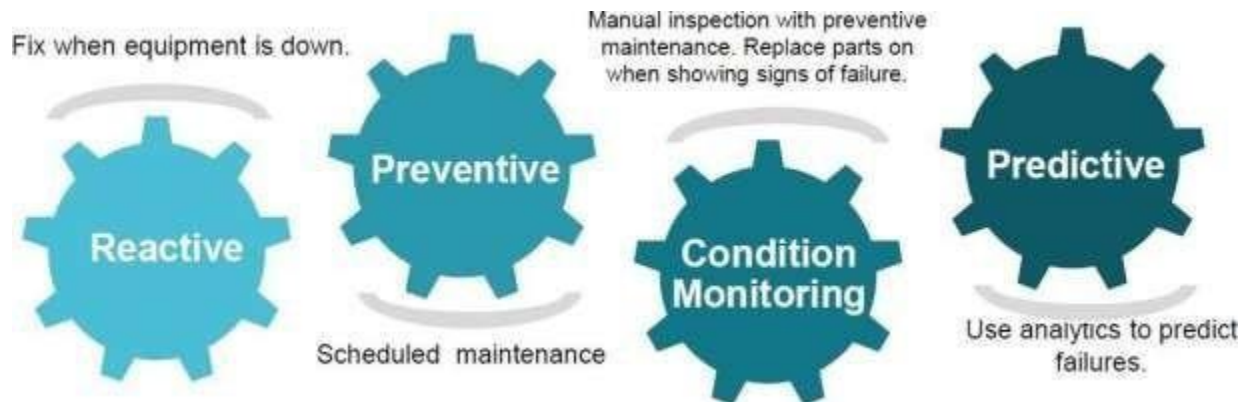
### iii. based Solution

UCT is one of the early adopters of LoRaWAN technology and providing solution in Agritech, Smart cities, IndustrialMonitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutionsetc.

### 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 Uniconvergetechnologies 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

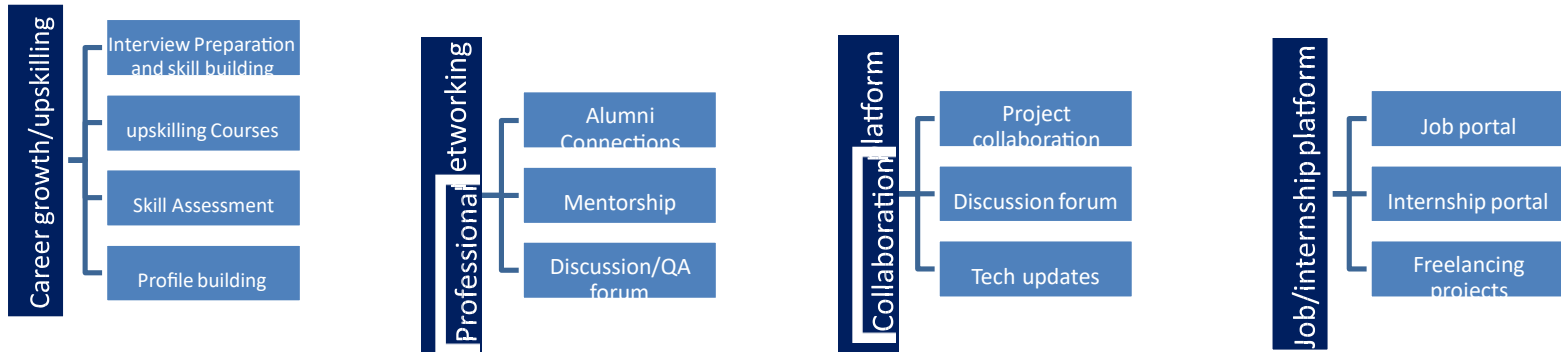


Seeing need of upskilling in self

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

- produce.csv
- datafile.csv

3 datafile (2).csv

4 <https://learn.upskillcampus.com/s/courses/6441224de4b0f11fbe0f621e/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.<sup>4</sup> 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

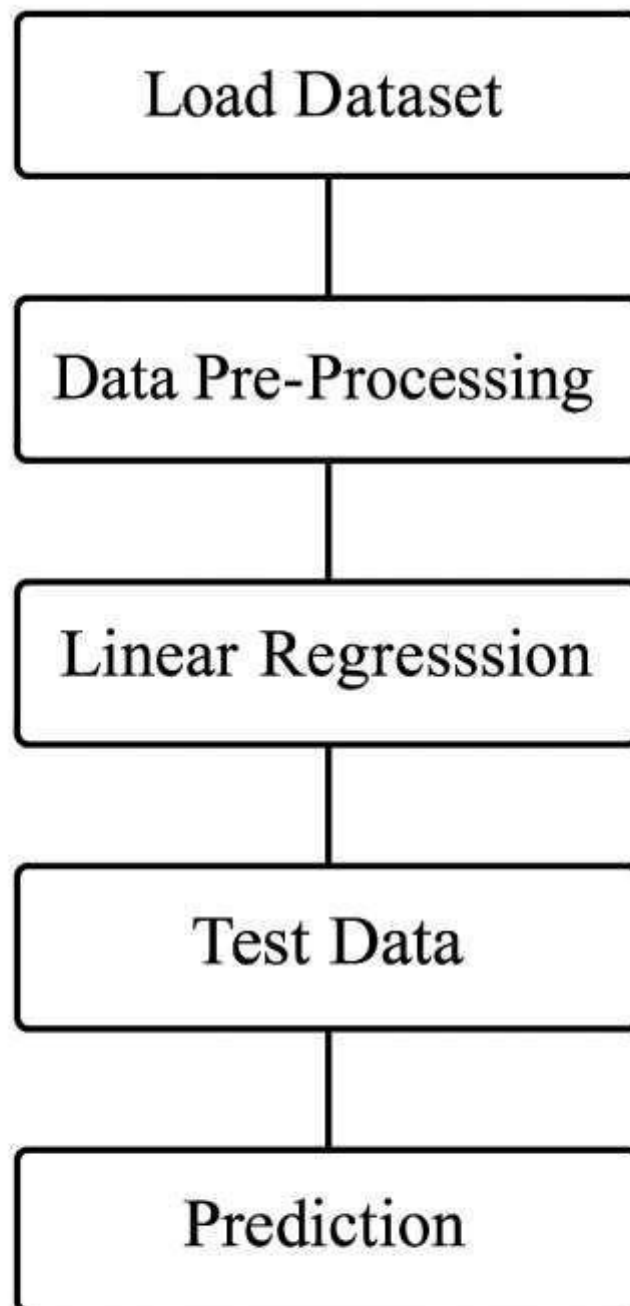
a new system which can predict the rate of production of crop yield accurately. In order to eradicate all such problems, we have proposed this new system, in which high yielding crop will be selected by considering most influencing parameters. This system helps the farmers to meet their crop yield production. 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.

**Code submission (Github Link)**

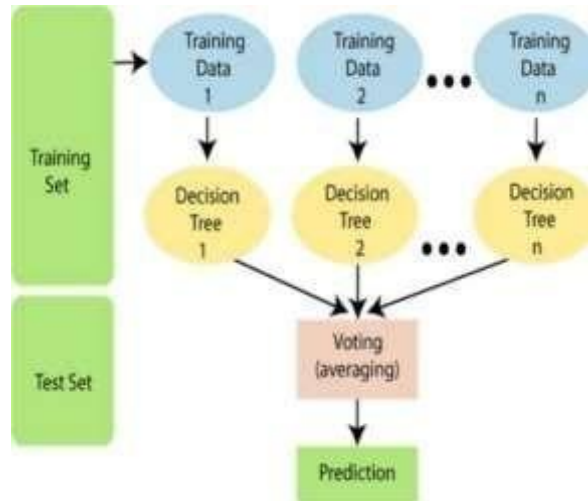
<https://github.com/shivareddy2003/upskillcampus/blob/main/AgricultureCropPrediction.py>  
[.ipynb](#)

**Report submission(Github Link)**

[https://github.com/shivareddy2003/upskillcampus/blob/main/AgricultureCropPrediction\\_Shiva\\_USC\\_UCT.pdf](https://github.com/shivareddy2003/upskillcampus/blob/main/AgricultureCropPrediction_Shiva_USC_UCT.pdf)

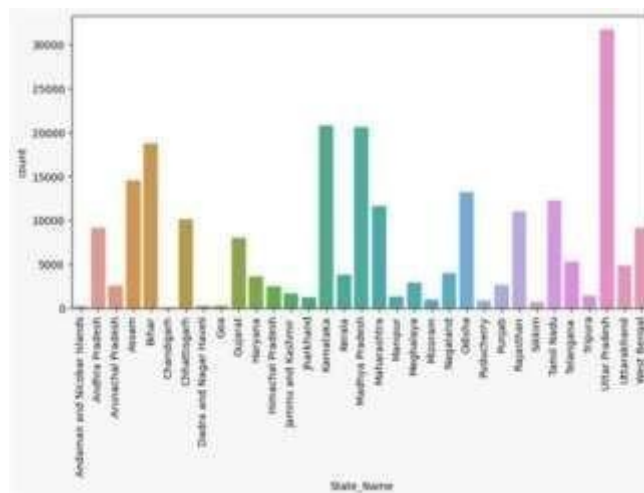


## 5 Proposed Design/ Model



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. 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

### 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.

### 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.

### 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. 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

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.