Industrial Internship Report on

”Prediction of Agriculture Crop Production in India”

Prepared by

Roshan Sayyad

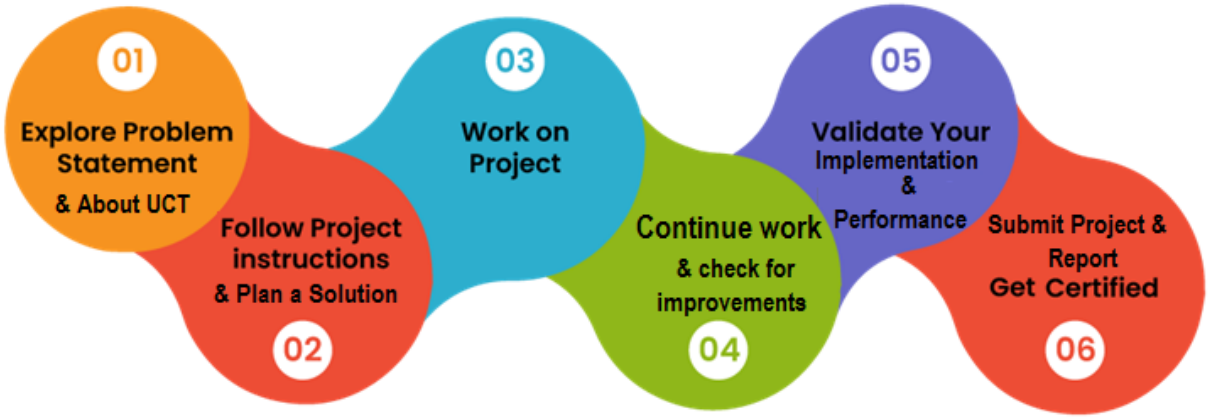
Executive Summary
This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT). This internship was focused on a project “Production of Agriculture Crop Production in India”, completed over a six-week period. This internship offered a valuable opportunity to gain exposure to industrial problems and design solutions using data science and machine learning techniques. The experience has been instrumental in my professional development.

TABLE OF CONTENTS

1 Preface

Summary of the whole 6 weeks' work.

In week 1, I embarked on my Data Science and Machine Learning internship by gaining and understanding of the company's mission, values and organizational structure. This week set the stage for the internship, providing a comprehensive overview of ongoing projects and tools and techniques essential for my role. Week 2 was dedicated to enhancing my knowledge through a detailed book on Data Science and Machine Learning. In week 3, the focus shifted to Probability and Statistics, crucial for data analysis. Week 4 introduced advanced machine learning concepts. I attended an introductory session and studied a detailed book on Machine Learning. A quiz helped reinforce these concepts. During week 5, the emphasis was on professional development. The company provided lecture on essential soft skills, effective interview techniques and strategies for both on-campus and off-campus placements. In week 6, the final week, I focused on preparing my project, "Prediction of Agriculture Crop Production in India", for submission.



This internship was made possible by the support and guidance of Upskill Campus, UCT and the IoT Academy. The program was meticulously planned to ensure comprehensive learning and practical application of data science and machine learning techniques.

I am grateful to all who have helped me directly or indirectly, including my mentors and peers. This internship has been a significant step in my career, equipping me with essential skills and knowledge.

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.

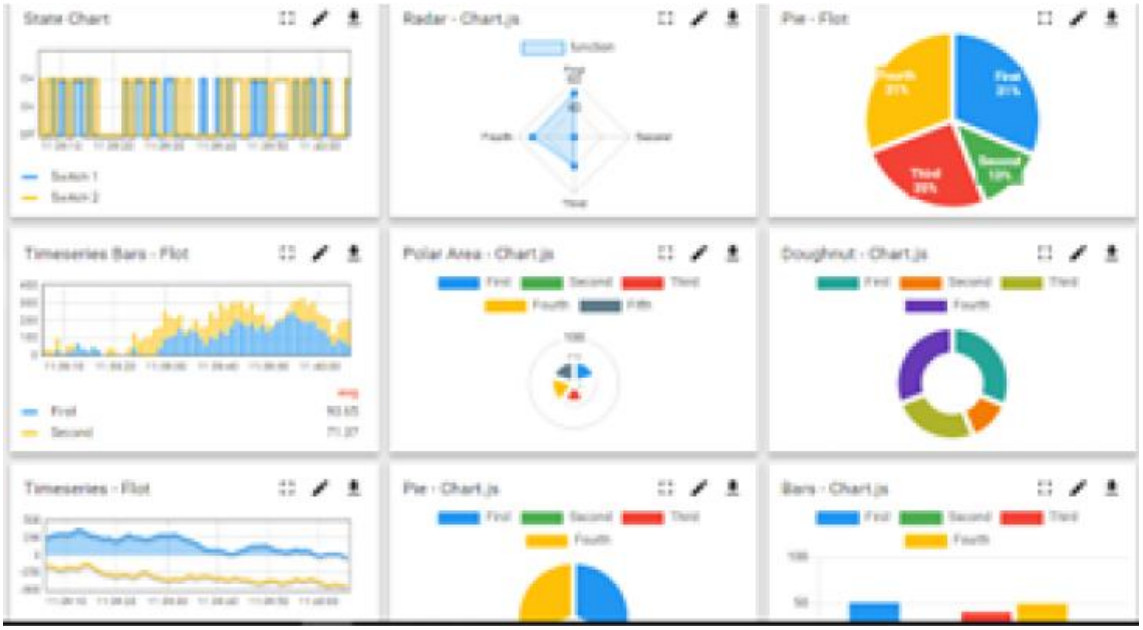
i. UCT IoT Platform (UCT Insight)

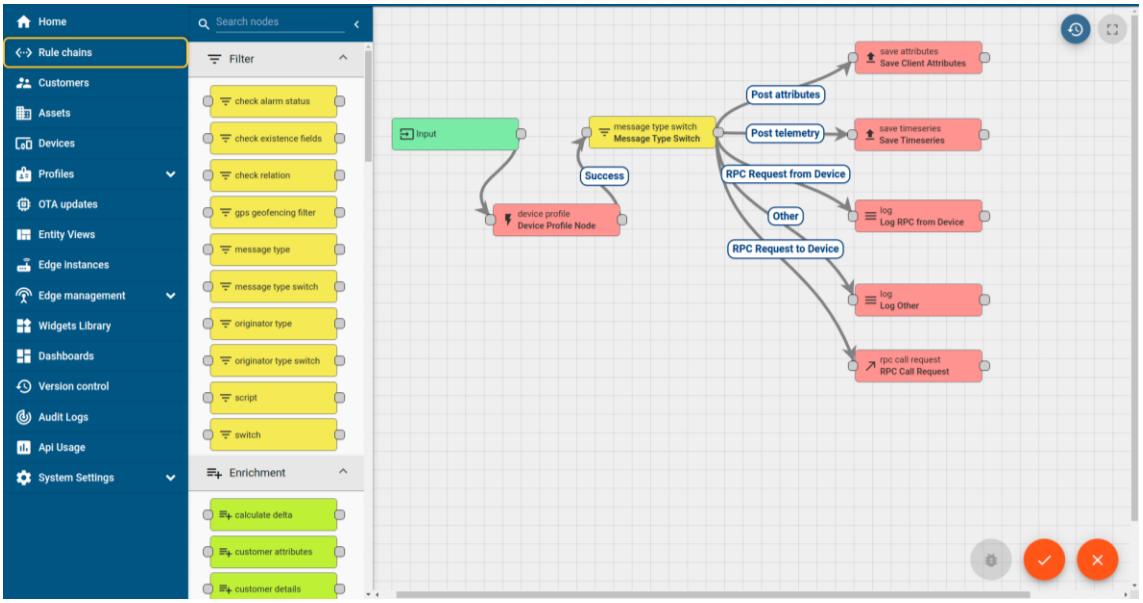
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- 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 (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.

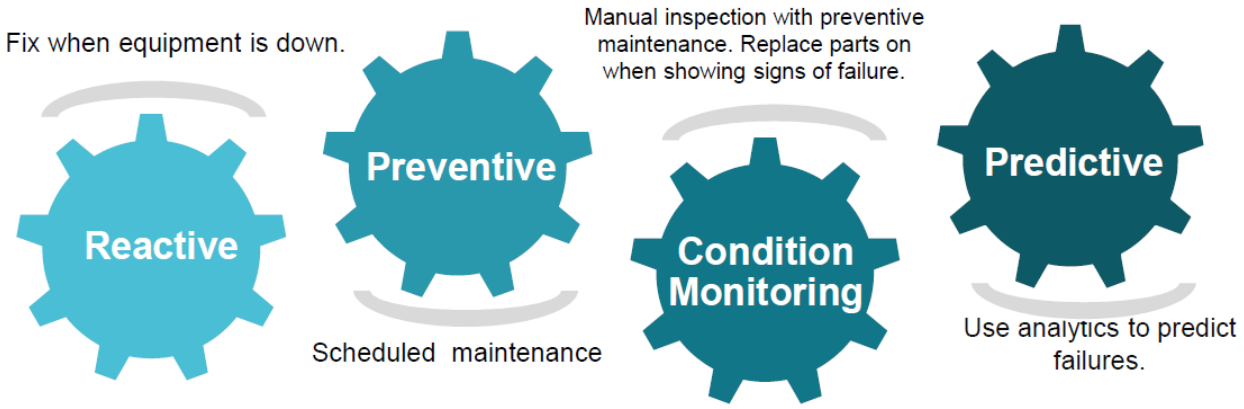
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1	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

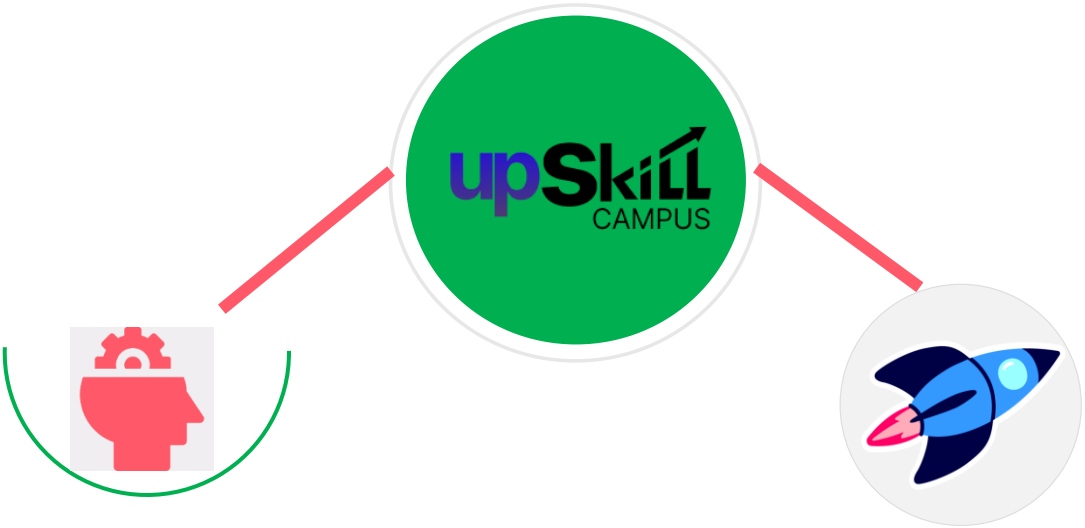
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.

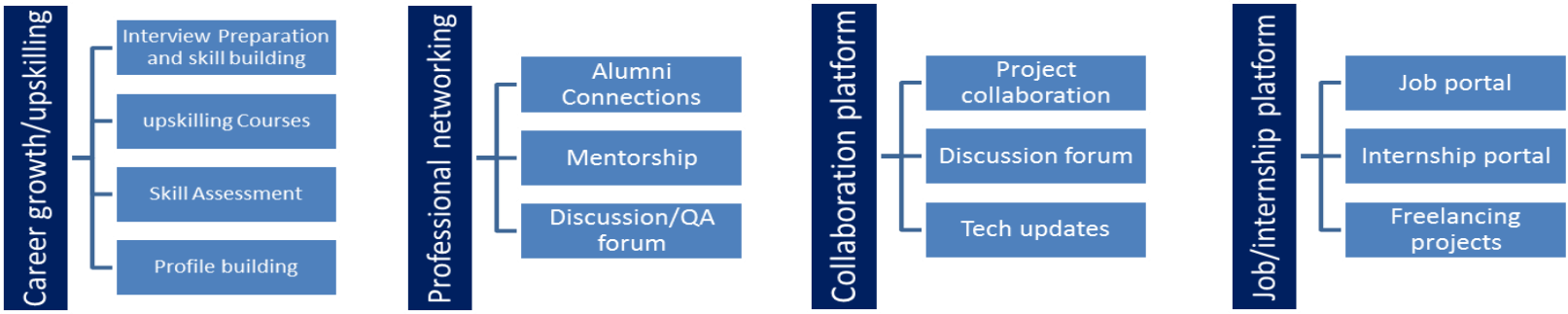
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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/



2.3 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.

2.4 Reference

[1] produce.csv

[2] datafile.csv

[3] datafile (2).csv

[4] https://learn.upskillcampus.com/s/courses/6441224de4b0f11fbe0f621e/take

3 Problem Statement

In the assigned problem statement

The project aims to predict agriculture crop production in India using data science and machine learning techniques. Accurate crop production forecasts can optimize resource allocation, improve food security and support farmer's livelihood.

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.

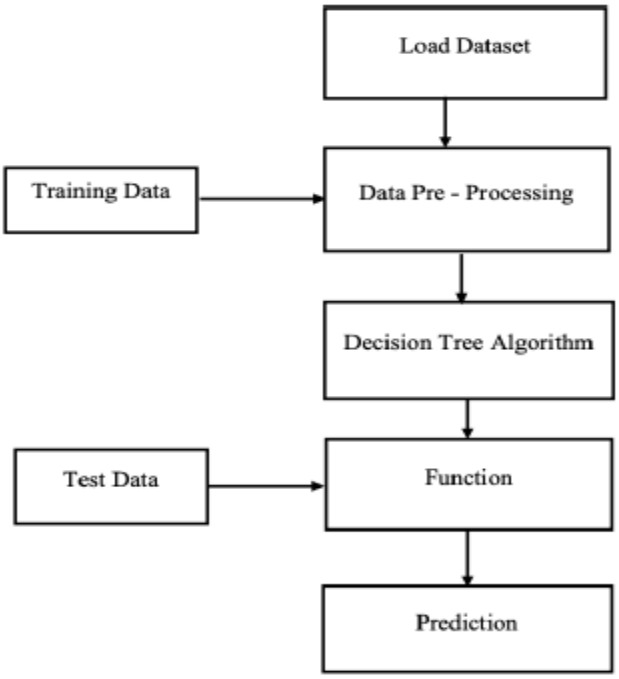
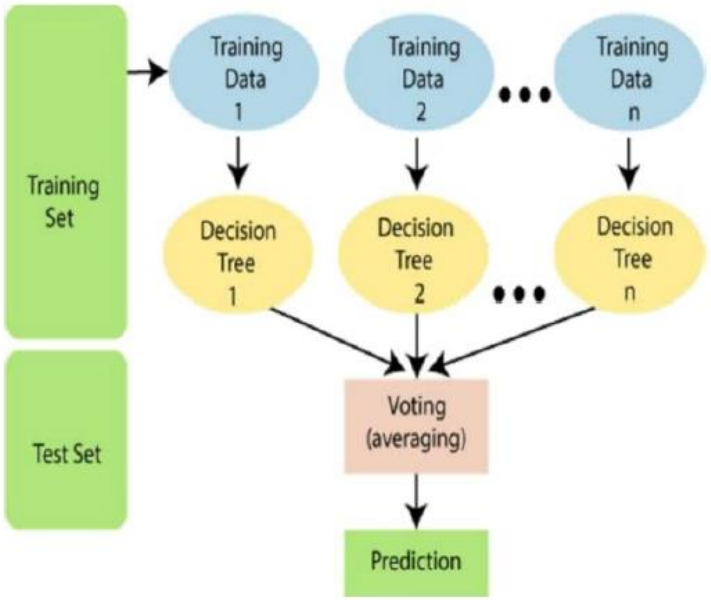


Fig.No.1: Proposed System

4.1 Code submission (Github link)

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.

5.1 High Level Diagram

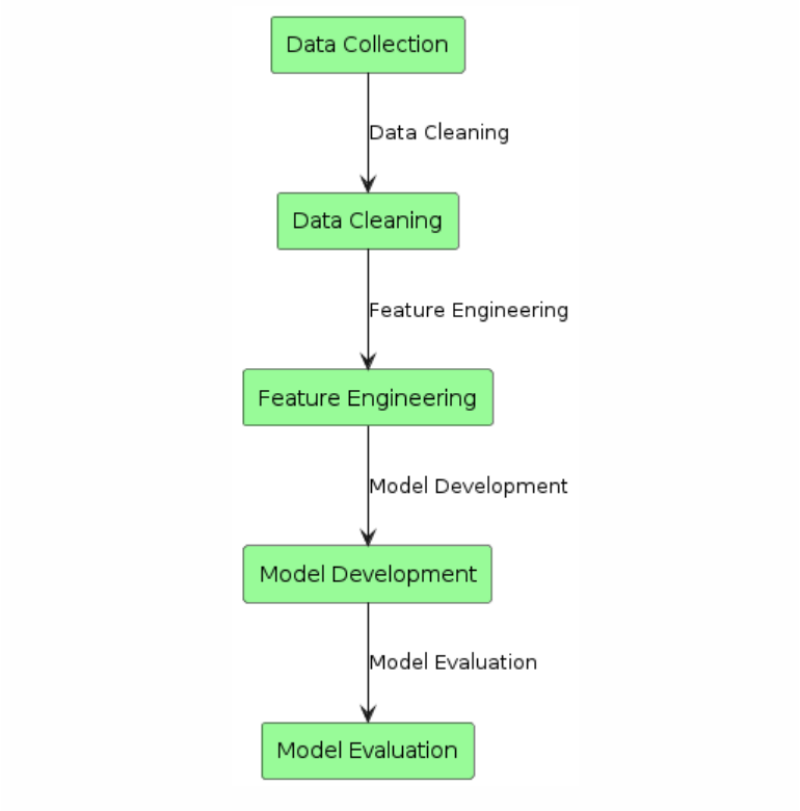


Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

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

This project has enhanced my understanding of data preprocessing, feature engineering, model development, and evaluation. The practical experience of handling real-world data issues and implementing machine learning algorithms has been invaluable for my career growth. 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.