Industrial Internship Report on

” Prediction of Agriculture Crop Production in India”

Prepared by

SHAHZIL SHAHID

Executive Summary
This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy incollaboration 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 projectincluding the report in 6 weeks’ time. My project was (Tell about ur Project) 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

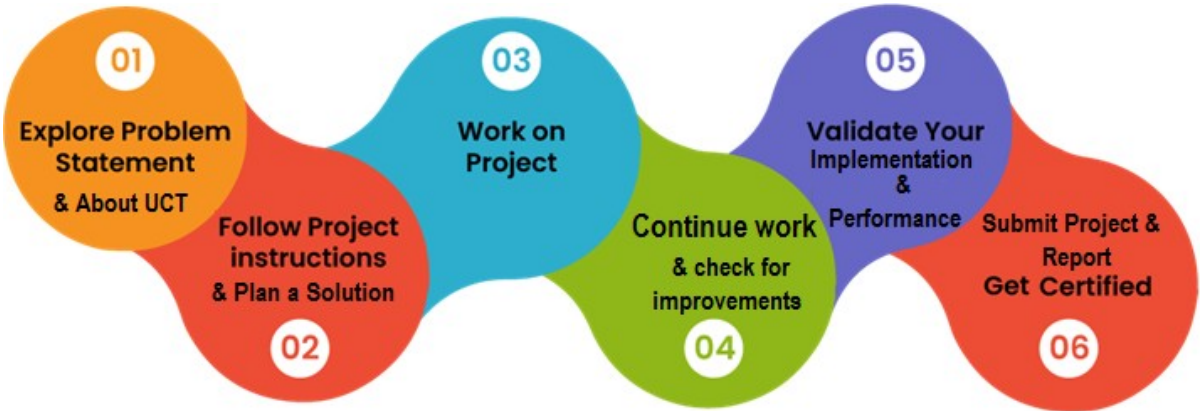
1 Preface

The data science and machine learning internship was of 6 weeks duration. During these weeks I worked on machine learning project. The topic chosen by me was 'Prediction of crop production in India'. During the first week, we were supposed to learn about the internship providing company and explore the various problem statements. During the 2nd week, we were expected to start working on the project. I checked various algorithms that could be used in order to design the solution. The idea was started to be implemented by the 3rd week. In the 4th week, the implementation was continued. In the 5th week, the quality and performance of the model was to be tested and any improvements to the project were to be added. The complete and final report is submitted in the 6th week.

Internships are very important for learning and growing individuals like us. They provide us with practical, real- world experience of working on real-life problems. Internships allow us to develop and hone our skills, explore different aspects of a field and develop qualities like adaptability and problem-solving skills.

India is a largely agriculture-based country. A lot of people's livelihood and the country's food needs are dependent on agriculture. Modernization of farming practices has helped us come a long way in providing for the needs. The crop yield or production majorly depends on the weather conditions, environmental changes, rainfall (which at times is uncertain), water management, and the utilization of pesticides. Therefore, farmers are not able accomplish expected yield of crop. Now a days data mining, machine learning as well as deep learning approaches are used by various researchers to enhance and improve the yield of crop and their quality. Machine Learning can gain proficiency with the machine without characterized computer programming, so it improves machine execution by distinguishing and portraying the consistency and pattern of drive information. In this work various machine learning approaches such as Linear Regression, Gradient Boosting Regressor, Random Forest Regressor, Decision Tree Regressor, Polynomial Regression, Ridge Regression have been used for yield prediction on crop yield dataset of different states and considering varied crops

I would like to thank Upskill campus for this opportunity as it helped me gain a lot of practical, hands-on experience on creating solutions to real world problems. It has been a really great experience working on this project.


01
Explore Problem Statement
& About UCT
02
Follow Project instructions
& Plan a Solution
03
Work on Project
04
Continue work
& check for improvements
05
Validate Your Implementation & Performance
06
Submit Project & Report
Get Certified

2 Introduction

2.1 About UniConverge Technologies Pvt Ltd

A company established in 2013 and working in Digital Transformation domain and providing Industrialsolutions with prime focus on sustainability and RoI.

For developing its products and solutions it is leveraging various **Cutting Edge Technologies e.g. Internetof Things (IoT), Cyber Security, Cloud computing (AWS, Azure), Machine Learning, Communication Technologies (4G/5G/LoRaWAN), Java Full Stack, Python, Front end** etc.

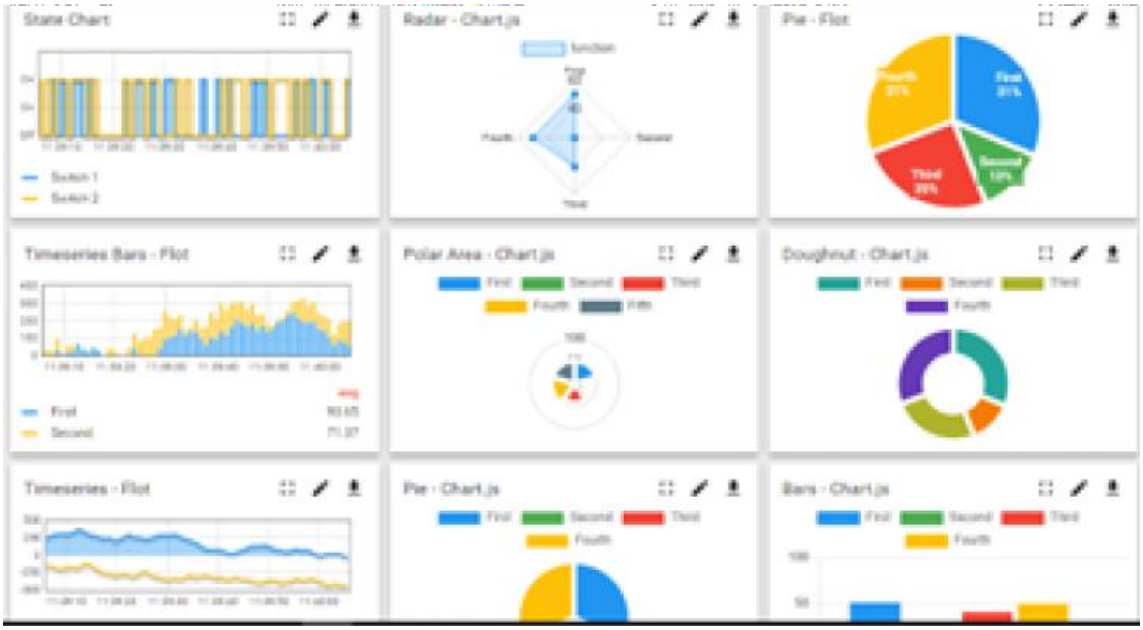
i. UCT IoT Platform (UCT Insight)

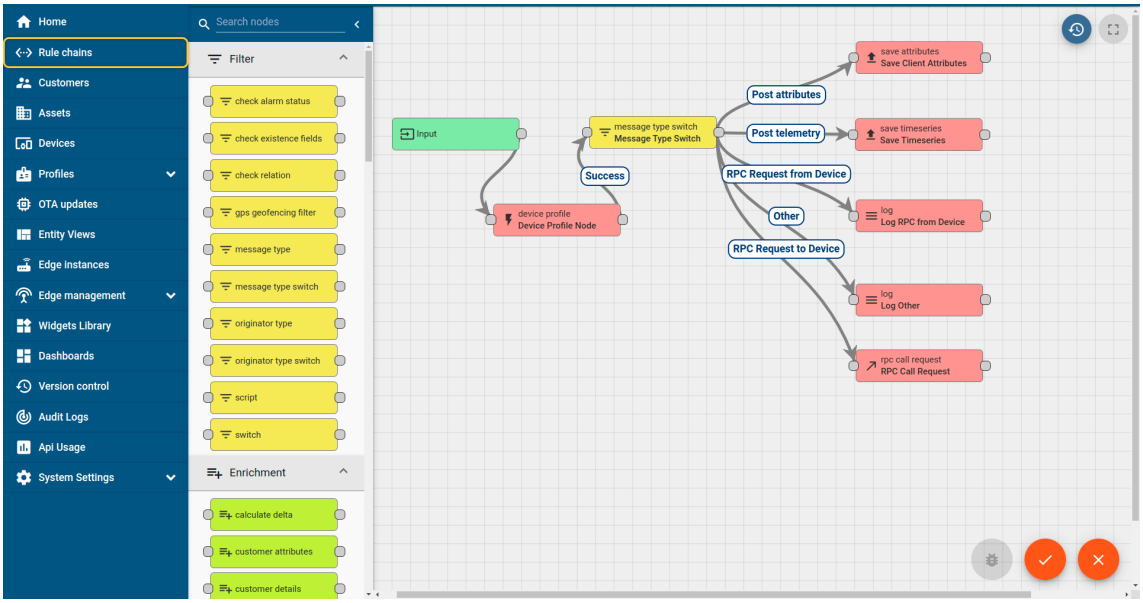
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable “insight” for your process/business. It has been built in Java for backend and ReactJSfor Front end. It has support for MySQL and various NoSql Databases.

- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, ModbusTCP, 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 toidentify the KPIs and also improve them.
- A modular architecture that allows users to choose the service that they what to start and thencan scale to more complex solutions as per their demands.

Its unique SaaS model helps users to save time, cost and money.

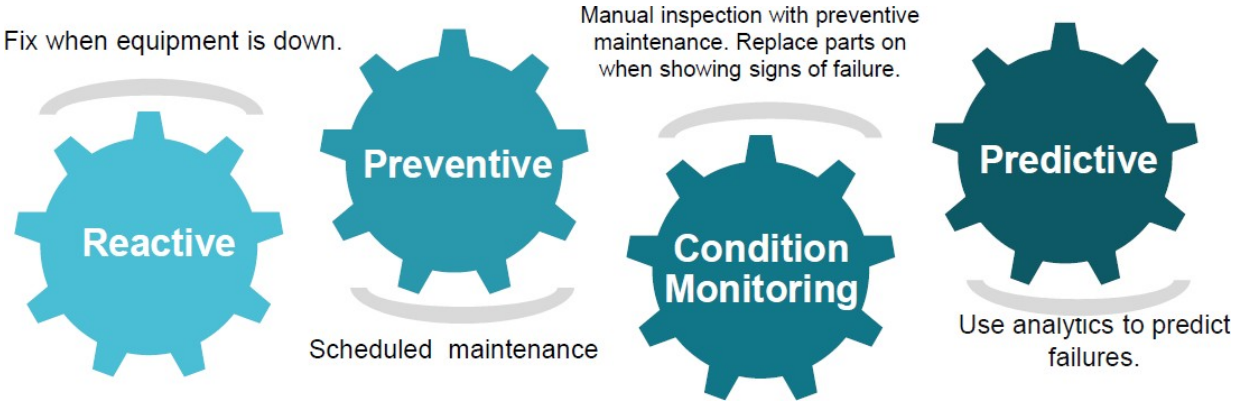
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1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

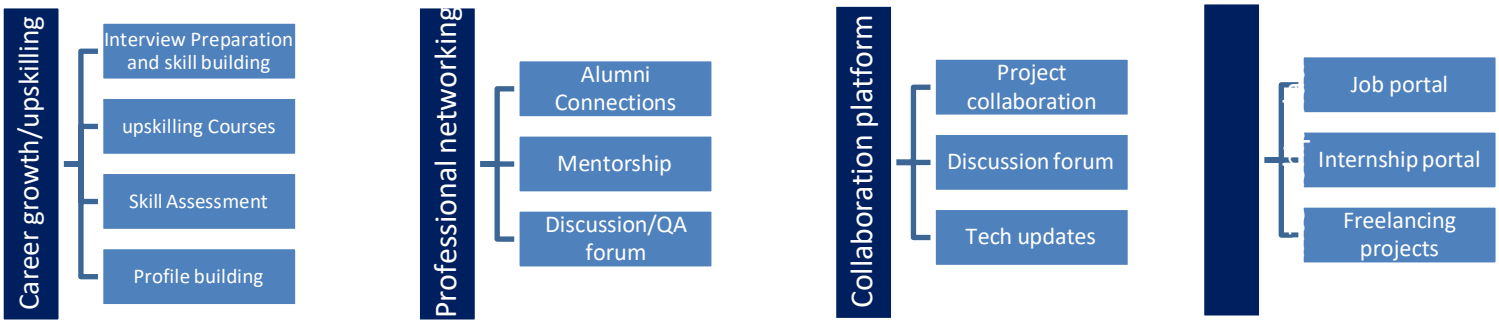
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveragingEmbedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.


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

2.3 The IoT Academy

The IoT academy is EdTech Division of UCT that is running long executive certification programs incollaboration 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] https://learn.upskillcampus.com/s/courses/64b90569e4b0203e0723aa96/take

3 Problem Statement

Agriculture is an important source of livelihood for farmers in India. Crop failure is a big issue due to inconsistency of weather conditions, rainfall, etc. There is a need to improve the sustainability of agriculture. There is a need of a system which can predict the yield of crop. This will help farmers take precaution to save the crop. The problem statement revolves around prediction of yield of crops considering different climatic conditions of India including various attributes. The aim of this project is to help farmers choose a better crop by predicting the yield in advance and take precautions.

4 Existing and Proposed solution

Existing Solution

Determining crop yields without relying on machine learning can still be achieved through variousTraditional and non-technological approaches. These methods include:

1. Crop estimation surveys – Conduct surveys with trained agronomists or extension agents toestimate crop yields.
2. Remote sensing and Satellite imagery - Utilize satellite or aerial imagery to monitor crop health,density, and growth throughout the growing season.
3. Historical and trend analysis - Analyze historical yield data for the specific crop and region.
4. Expert opinion and farmer surveys
5. Soil testing and analysis

Limitations

Traditional methods of crop yield prediction have several limitations that make them less accurate. Theselimitations are:

1. Limited accuracy
2. Time – consuming
3. Reliance on Subjective data
4. Limited spatial coverage
5. Dependence on historical data
6. Data limitations
7. Limited scalability

Proposed Solution

There is a critical need to form a model that can give accurate predictions about the crop yield. In the proposed model, high yielding crop will be selected by considering the most influencing parameters. Thismodel helps farmers predict the production of crop and thus reduces the chances of failure of crop. The prediction model is created using the machine learning techniques such as Random Forest Regressor andDecision tree regressor. The input parameters given are state name, area, crop and season.

Value Addition

Machine learning-based crop yield prediction models offer several significant value additions comparedto traditional methods. These advancements are primarily driven by the ability of machine learningalgorithms to analyze large datasets, learn complex patterns, and make accurate predictions. Some keyvalue additions are as follows:

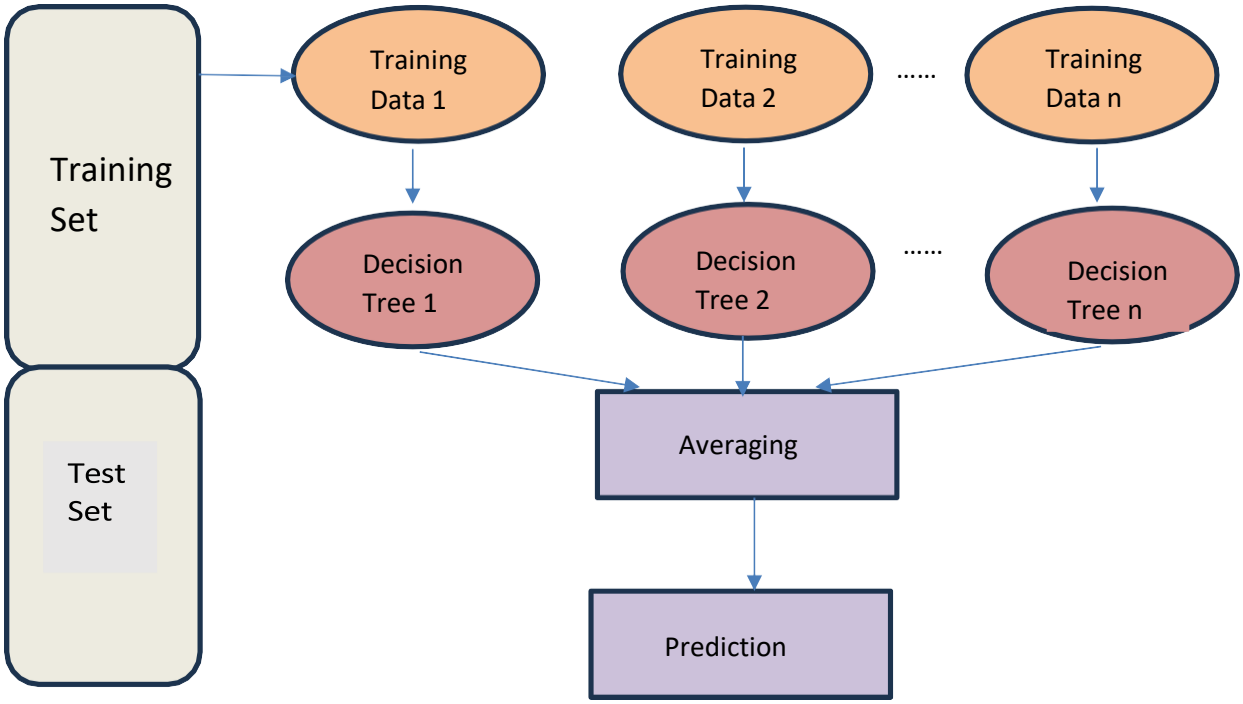
1. Improved accuracy: Machine learning models can provide more accurate crop yield predictionsby analyzing historical data, current environmental conditions, and other relevant factors. Theycan capture intricate relationships and nonlinear patterns that traditional methods may miss, leading to more precise predictions.
2. Timely predictions: Machine learning models can provide real-time or near-real-time crop yieldestimates during the growing season which allows the farmers and stakeholders to make informed decisions about irrigation, pest control, and harvesting schedules.
3. Scalability: The scalability enables agricultural organizations to make strategic decisions andallocate resources effectively.
4. Adaptability: Machine learning models can adapt to changing conditions incorporate new data asit becomes available.
5. Early warning systems: Machine learning models can serve as early warning systems for potentialcrop failures or reduced yields due to adverse weather conditions, pests, or diseases allowing farmers to take proactive measures to mitigate losses.
6. Risk management: Machine learning models can assess and quantify risks associated with cropproduction, helping farmers make informed decisions about risk mitigation strategies.

4.1 Code submission (Git-hub link)

Repository link: https://github.com/Shahzilshahid/SHAHZIL-SHAHID/tree/main

5 Proposed Design/ Model

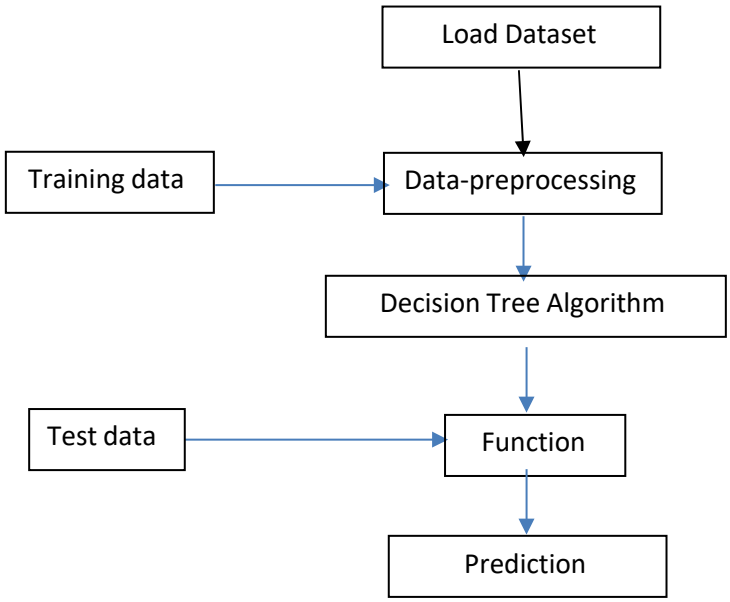
Flowchart of crop yield prediction:



In the proposed system, machine learning techniques are applied to predict the rate of crop production.The programming language used is Python as it is a widely used language in data science and machine learning applications. First the collected data set will be uploaded. Machine learning techniques used in this project are Random Forest Regressor and Decision Tree Regressor. The results of the prediction model depend directly on the accuracy of information in the collected datasets. More accurate information in turn gives accurate predictions. Hence data pre-processing is an important aspect of the process. After the function is created by implementing ML algorithm, some test data is given to it to testthe performance of the system. Hence, by improving we finally create a model that can successfully predict the required thing.

5.1 High Level Diagram (if applicable)

Diagram of the process followed to build 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.

Crop Yield Prediction: This module returns the predicted production of crops based on the user's input. Ifthe 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 innext 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 aregiven 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 valuesare 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 internship has been a really great opportunity to learn about data science and machine learning andwork on a real-world problem. First, I studied about the internship providing company, UCT, about their domains of work and technologies used. Read and understood the instructions about the internship and the process. I started by learning about the basics of data science and its applications. I referred to the video content on the upskill campus profile to learn about machine learning in depth. Went through all the project topics mentioned. Selected the project for the internship in project category A: Agriculture- Project 4- Prediction of Agriculture Crop Production in India. Went through some articles to understand the problems faced by farmers in India, the measures taken, the traditional approaches used to predict the crop yield and their accuracy. This helped me understand the problem statement better. Next, I went through the data provided about the previous years' crop production. Read the e-book 'Introducing Datascience and Machine learning' to understand the concepts better. Glanced over the article 'Impact of Bigdata on business', understood what big data means, where it is useful. Learnt about the difference
between a data scientist and a data analyst. Read another blog on 'AI vs Data Science'. It helped me understand the difference between the two terms and how they are correlated. It gave me insight intomapping career in either field and the relevant skills required. AI and data science are
interconnected. So, specialization in either field requires a working knowledge of the other. To begin using ML-AI, studying data analysis is more necessary than understanding data science. No business or firm would be able to run without data science . During the tenure of this internship, I studied about various machine learning algorithms, how they are implemented, how to choose the ideal algorithm, how the results vary depending on data.
Also, attempted quizzes on the platform to test my knowledge. Learnt about the essential skills to become a data scientist, which includes programming, Statistics and Mathematics, Machine learning, Deep Learning, and Big data.

These learnings helped me advance my existing skillset and allowed to dive deeper in the field of data science. This would surely help me in my future work in the same field, has helped me enhance my resume by adding an innovative project to it.

8 Future work scope

This system is proposed to help the farmers and thus allow them to grow financially. The recommendation system helps the farmers predict the crop yield given the data. It would in turnbenefit them since now they would know which crop would give the best yield. The user needs to give input data such as soil type and area. The datasets were collected, studied and trained using ML tools. The system would track the user location and fetch needed info from the backend based on the location. Thus, appropriate results will be provided. Moreover, the systemalso tells the user the right time to use the fertilizer.

Another feature that I wanted to add to this system was the prediction of price of crop accordingto the present market conditions. However, given the time constraint I could only work on this much. Hope the report explains everything well.

Thank you!