Industrial Internship Report on

"Prediction of Agriculture Crop Production in India"

Prepared by

1. **Sandip Rathwa**
2. **Sunil Patel**

| *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/problem statement provided by UCT. We had to finish the project including 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

In the sixth week of our internship, we learned about project skills and how to write an industrial report. We also focused on our project, which is to predict agricultural crop production in India. We performed various operations on the dataset to help us with our data analysis.

An industrial report is a formal document that is used to communicate the findings of a research project or technical investigation. It is typically written for a professional audience, such as engineers, scientists, or managers. When writing an industrial report, it is important to be clear, concise, and accurate. You should also use appropriate technical language and avoid using jargon. The report should be well-organized and easy to read. It should also be properly formatted and cited.
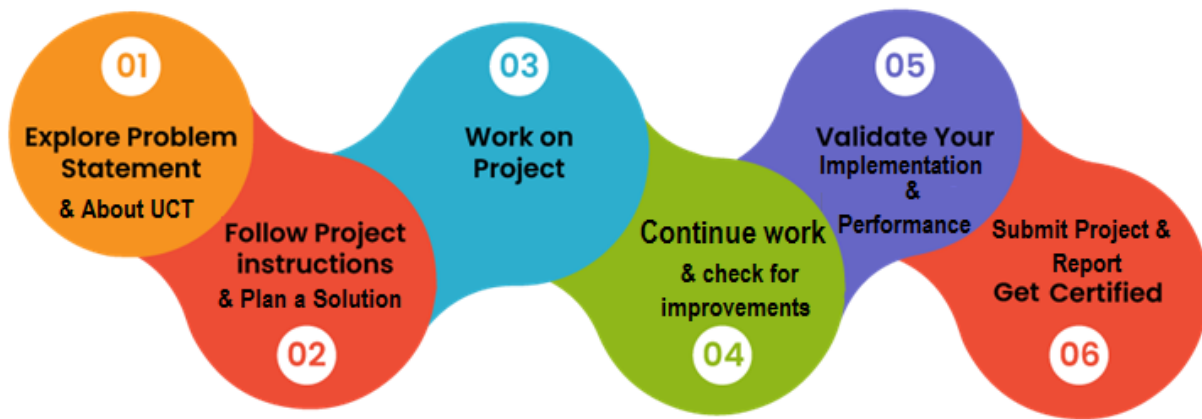
I had the chance to work with seasoned professionals and obtain insightful knowledge about the real-world uses of principles from the discipline of data science throughout my time there. I was able to use my knowledge and abilities in practical situations by taking on hard projects, which helped me develop my problem-solving skills and push the limits of my data scientist abilities. The presence of seasoned mentors and encouraging coworkers inside the organisation produced a nurturing environment where I was able to obtain direction, feedback, and helpful advice to advance my career as a data scientist. The presence of seasoned mentors and encouraging coworkers inside the organisation produced a nurturing environment where I was able to obtain direction, feedback, and helpful advice to advance my career as a data scientist.

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
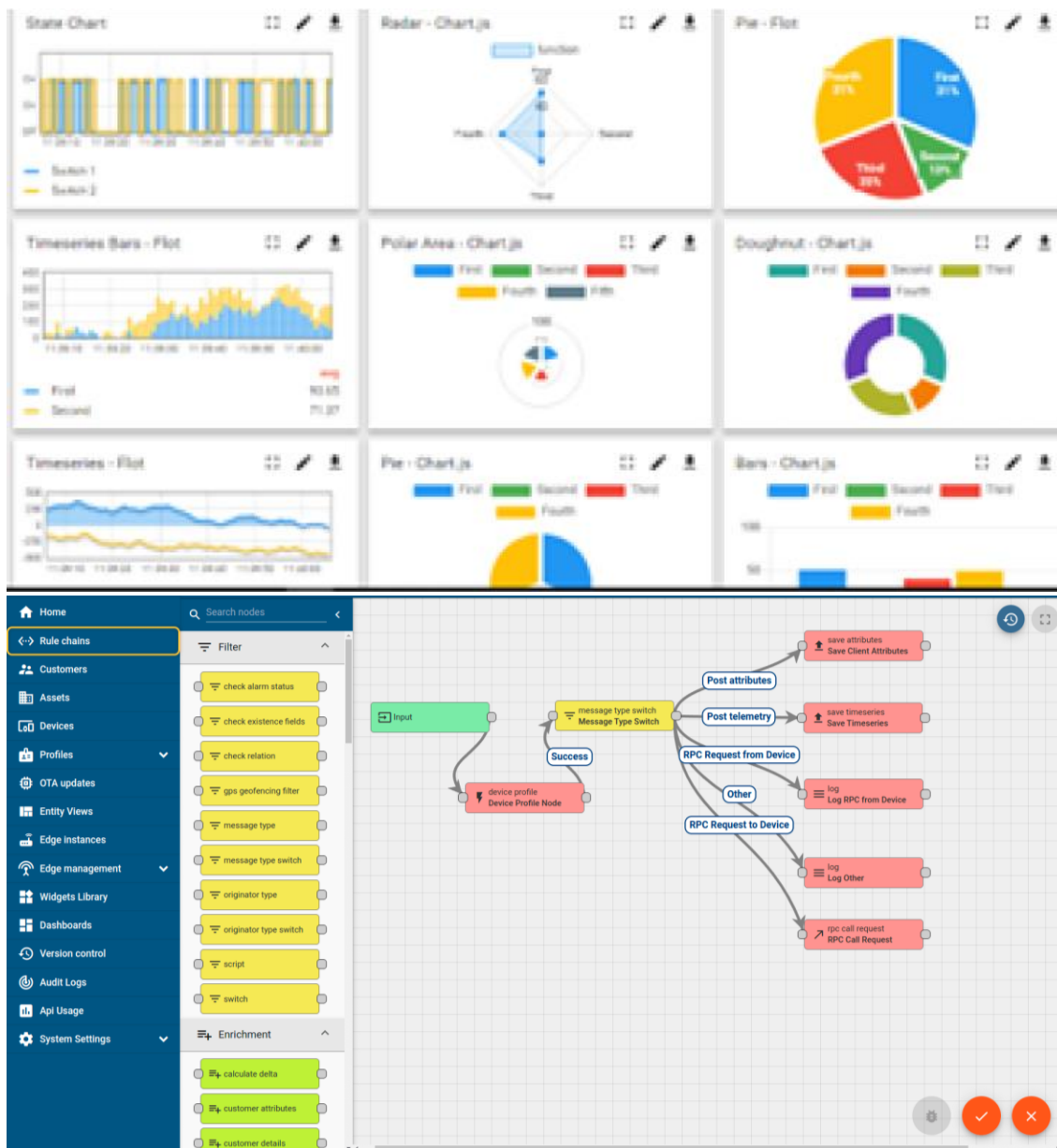
ii. Smart Factory Platform (FACT◯RY WATCH)

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.
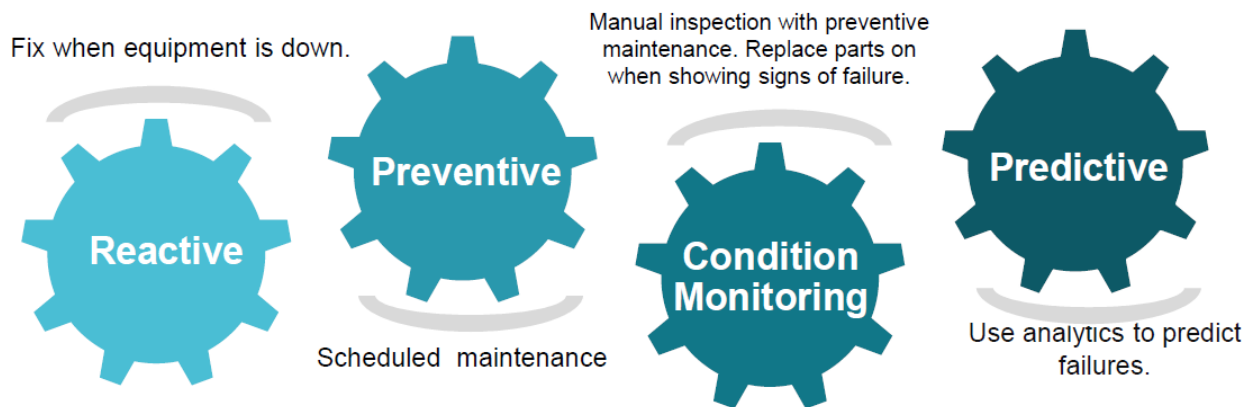
2.2   About upskill Campus (USC)

upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.
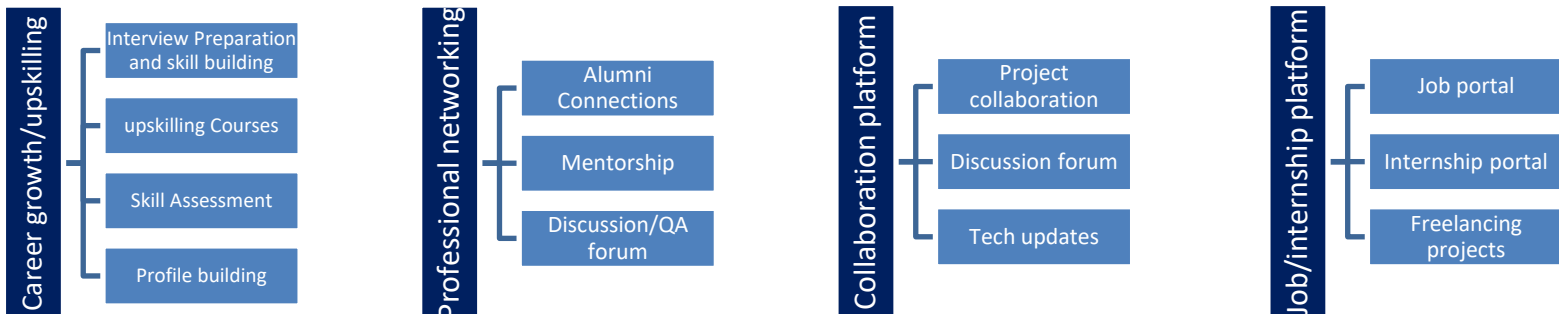
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.
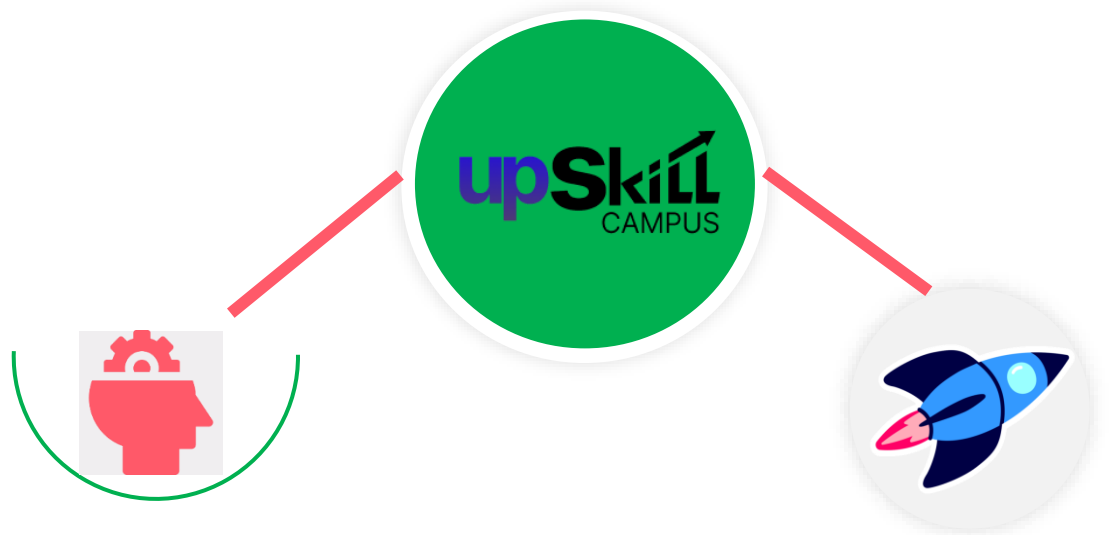
Seeing need of upskilling in self paced manner along-with additional support services e.g. Internship, projects, interaction with Industry experts, Career growth Services

upSkill Campus aiming to upskill 1 million learners in next 5 year

https://www.upskillcampus.com/

Career growth/upskilling
- 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/internship platform
- 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 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]   "Python for Data Analysis" by Wes McKinney

[2]   "Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" by AurélienGéron

[3]   "Data Science for Business" by Foster Provost and Tom Fawcett

[4]   "The Elements of Statistical Learning" by Trevor Hastie, Robert Tibshirani, and Jerome Friedman

[5]   "Data Science from Scratch" by Joel Grus

[6]   Kaggle (www.kaggle.com): A platform for data science competitions and datasets with a wealth of resources and notebooks from the community.

[7]   Towards Data Science (towardsdatascience.com): An online publication with a vast collection of articles covering various data science topics.

2.6 Glossary

Terms	Acronym
Machine learning	Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate in predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.
Artificial Intelligence	Artificial intelligence (AI) is the ability of a machine or computer program to simulate human intelligence. AI systems are able to learn from data and experience, and they can adapt to new situations.
Data Scientist	A data scientist is a professional who uses statistics, computer science, and domain expertise to extract knowledge and insights from data. Data scientists work in a variety of industries, including technology, finance, healthcare, and retail.
Data Engineer	A data engineer is a software engineer who specializes in building and maintaining data infrastructure. Data engineers design, build, and maintain the systems and pipelines that allow organizations to collect, store, process, and analyze data.
BigData Analytics	Big data analytics is the process of collecting, examining, and analyzing large amounts of data to discover market trends, insights, and patterns that can help organizations make better decisions.

3   Problem Statement

Agriculture is the backbone of the Indian economy. Due to global warming and climate change, traditional farming in the regular months has been distorted, and crops have been ruined. This is the most common phrase seen today. This not only causes economic losses but is also the main reason for farmer suicide. Now that agriculture needs support, the time has come for technology to take over. For a crop to grow, favorable soil conditions, ambient rainfall, and temperature are necessary. So we have formulated an analysis based on the prediction of a favorable crop based on temperature, current rainfall, and soil conditions. Agriculture is the backbone of the Indian economy. Due to global warming and climate change, traditional farming in the regular months has been distorted, and crops have been ruined. This is the most common phrase seen today. This not only causes economic losses but is also the main reason for farmer suicide. Now that agriculture needs support, the time has come for technology to take over. For a crop to grow, favorable soil conditions, ambient rainfall, and temperature are necessary. So we have formulated an analysis based on the prediction of a favorable crop based on temperature, current rainfall, and soil conditions.

4 Existing and Proposed solution

- **Statistical models:** Statistical models are used to identify relationships between historical crop yields and other factors, such as weather, soil type, and fertiliser use. These models can then be used to predict future crop yields based on current conditions.
- **Machine learning models:** Machine learning models are trained on historical data to learn to predict crop yields. These models can be more accurate than statistical models, but they require more data to train.
- **Remote sensing:** Remote sensing data, such as satellite imagery, can be used to monitor crop growth and health. This data can be used to predict crop yields by identifying patterns and trends in the data.

4.1 Code submission (Github link)

https ://github. com/sandip2264/upskillcampus/blob/main/ Crop-Production in India.java

4.2 Report submission (Github link) :

https ://github. com/sandip2264/upskillcampus/blob/main/Crop-Production in India _Sandip_USC_UCT.pdf

5   Proposed Design/ Model

- **Define the problem.** What crop are we trying to predict the yield for? What region are you interested in? What factors are you going to use to predict the yield?
- **Collect data.** We will need to collect data on historical crop yields, weather, soil type, and other relevant factors. This data can be obtained from a variety of sources, such as government agencies, private companies, and research institutions.
- **Clean and prepare the data.** Once we have collected the data, you will need to clean it and prepare it for analysis. This may involve removing outliers, correcting errors, and converting the data into a consistent format.
- **Choose a prediction method.** There are a number of different methods that can be used to predict crop yields. Some of the most common methods include statistical models, machine learning models, and remote sensing.
- **Train the model.** If we are using a machine learning model, you will need to train it on historical data. This involves feeding the model the data and allowing it to learn the relationships between the different factors.
- **Evaluate the model.** Once the model is trained, we need to evaluate its performance on a held-out test set. This will give us an idea of how accurate the model is likely to be with the new data.
- **Deploy the model.**Once we are satisfied with the performance of the model, we can deploy it into production. This may involve integrating the model into a software application or making it available as a web service.

5.1 High Level Diagram (if applicable)

Agricultural Production Over Time
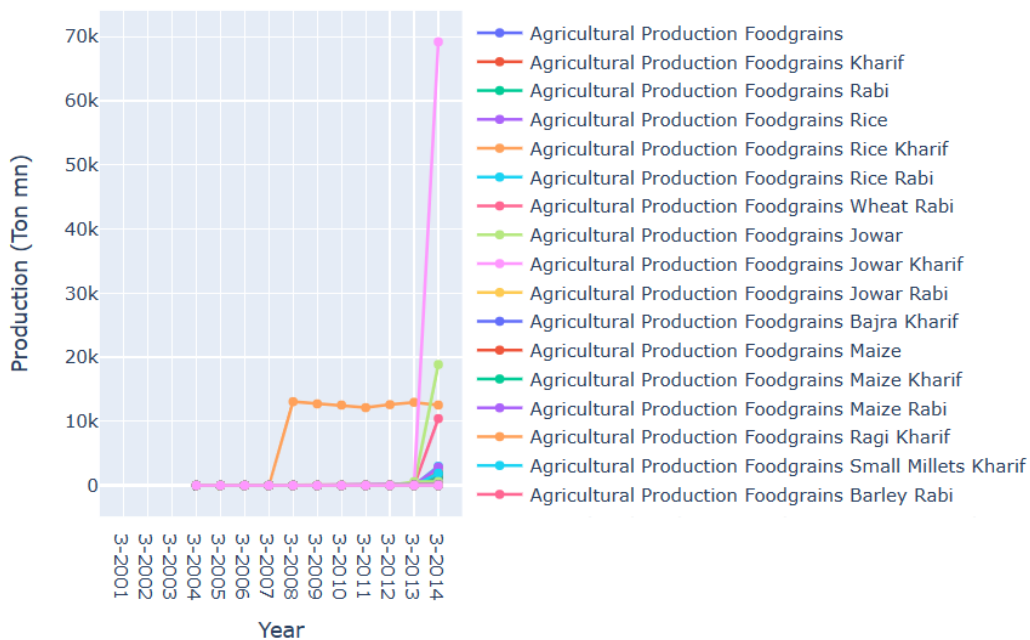
5.2 Low Level Diagram (if applicable)
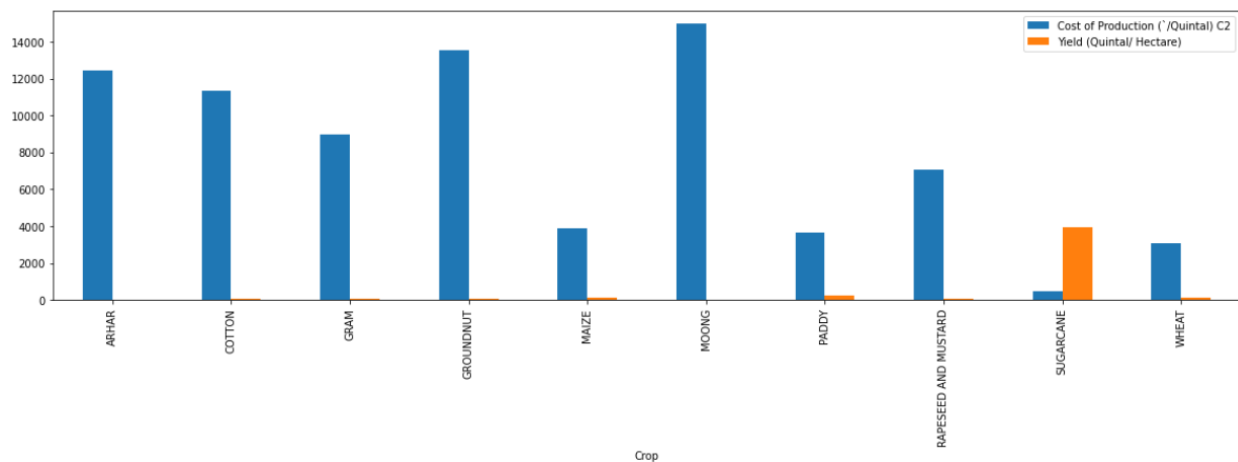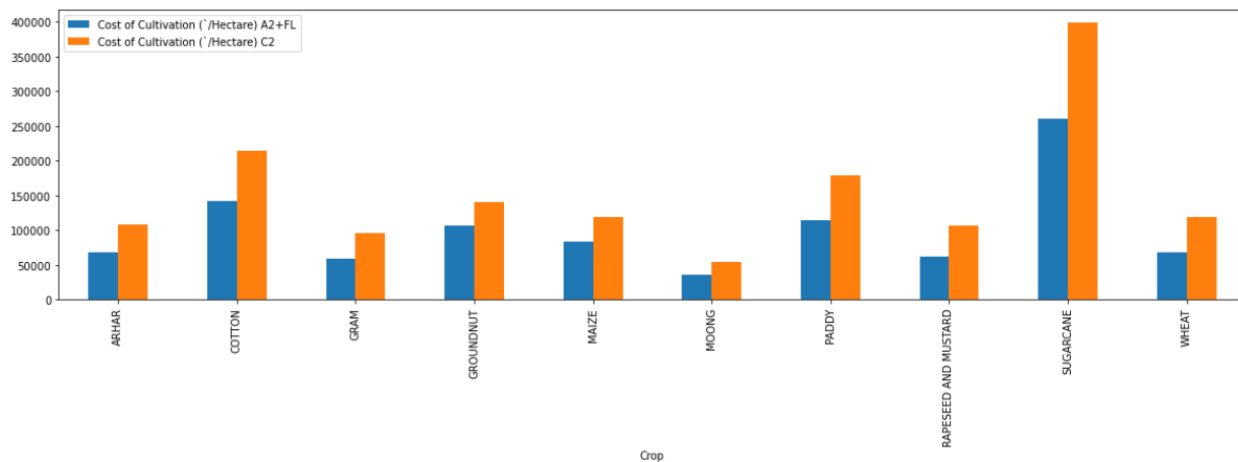
5.3 Interfaces (if applicable)

Start
Define the problem
Collect data
Clean and prepare the data
Choose a prediction method
Train the model
Evaluate the model
Deploy the model
End

- The flowchart shows the steps involved in predicting agricultural crop production in India. The first step is to define the problem, which includes identifying the crop and region of interest as well as the factors that will be used to predict the yield. Next, data on historical crop yields, weather, soil type, and other relevant factors is collected. The data is then cleaned and prepared for analysis, which may involve removing outliers, correcting errors, and converting the data into a consistent format.

6. My learnings

Enhanced Data Analysis abilities: As a result of my time working for the company, I have acquired advanced data analysis abilities, such as the capacity to draw insights from large, complicated datasets, spot patterns and trends, and come to data-driven conclusions.

Competency in Machine Learning: Working on a variety of machine learning projects allowed me to develop a thorough understanding of various algorithms, model selection strategies, and evaluation metrics. I gained the knowledge and abilities necessary to create, test, and use machine learning models for predictive and prescriptive analytics.

Data Preprocessing and Feature Engineering: I developed competence in dealing with real-world data issues, including data cleansing, addressing missing values, outliers, and executing feature engineering procedures to convert raw data into meaningful features that improve model performance.

Practical expertise in Model Deployment: I received first-hand knowledge of the challenges associated with integrating machine learning models with current systems and assuring scalability, reliability, and performance. I also gained practical expertise in deploying machine learning models into production environments.

Data visualisation and communication: I honed my skills in producing powerful data visualisations and successfully sharing insights with both technical and non-technical stakeholders. I gained knowledge on how to communicate intricate findings in a compelling and simple way to encourage data-driven decision-making.

cooperation and project management: By participating in cross-functional teams, I improved my cooperation and project management abilities. I gained expertise in coordinating project goals, controlling deadlines, and productively interacting with team members.

7. Future work scope

- **Develop more accurate and reliable prediction models.** This could involve using more data, more sophisticated machine learning algorithms, or new remote sensing technologies.
- **Make the prediction models more accessible to farmers.** This could involve developing mobile apps or web services that farmers can use to easily access the predictions.
- **Integrate the prediction models with other agricultural systems.** For example, the predictions could be used to develop decision support systems for farmers or inform agricultural insurance policies.
- **Develop prediction models for other crops and regions.** Currently, most crop yield prediction models are focused on a small number of crops and regions. There is a need to develop models for other crops and regions, especially in developing countries.
- **Investigate the impact of climate change on crop yields.** Climate change is a major threat to agriculture. It is important to develop prediction models that can account for the impact of climate change on crop yields.