Industrial Internship Report on

”Prediction of Agriculture Crop Production in India”

Prepared by

Ajinkya Sanjay Shitole

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 “Prediction of Agriculture Crop Production in India” 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

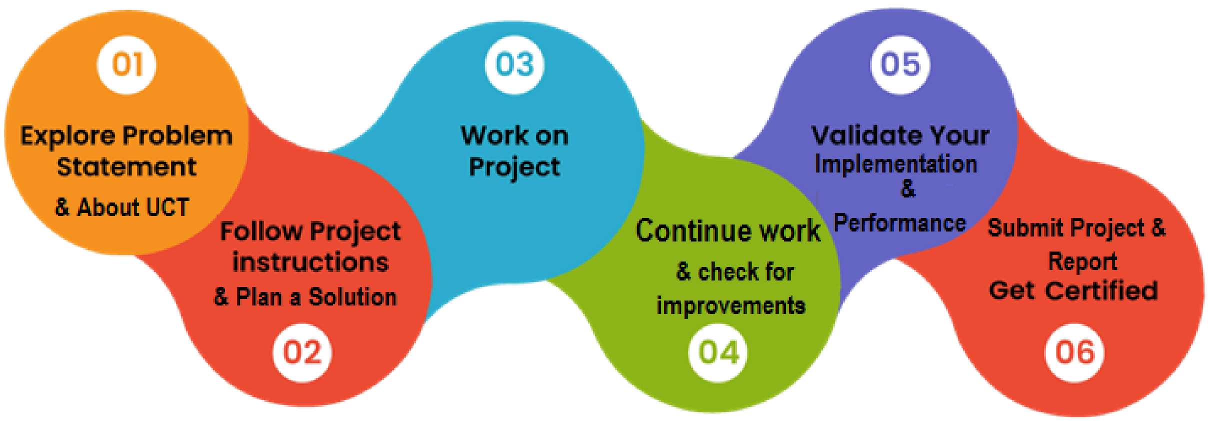
Summary of the whole 6 weeks' work.

About need of relevant Internship in career development.

Brief about Your project/problem statement.

Opportunity given by USC/UCT.

How Program was planned



Your Learnings and overall experience.

Thank to all (with names), who have helped you directly or indirectly.

Your message to your juniors and peers.

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.

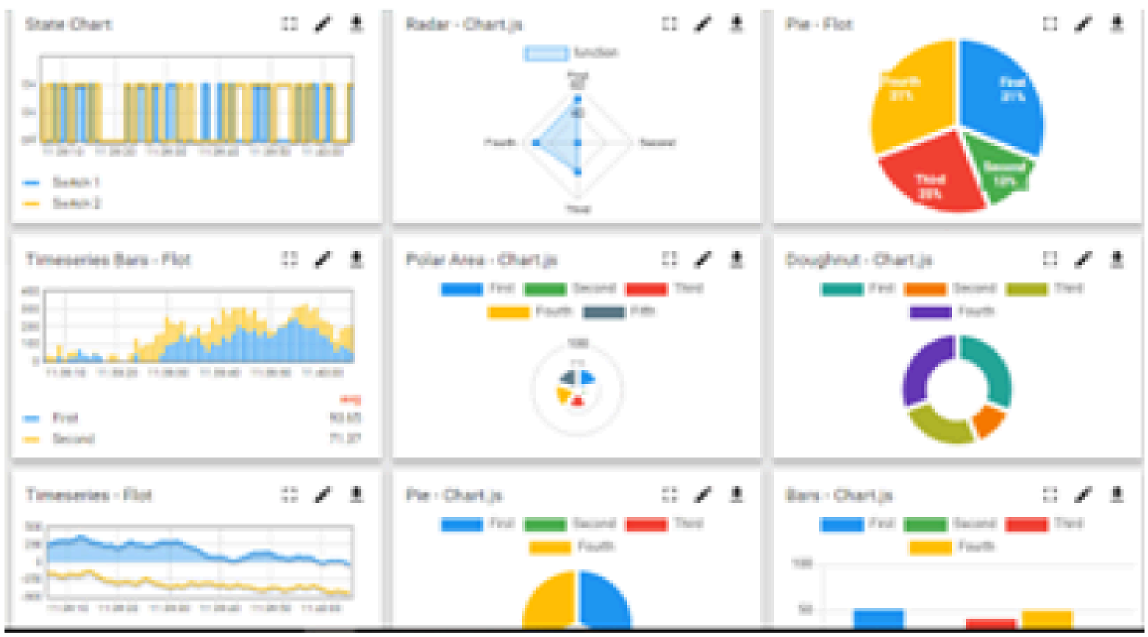
i. UCT IoT Platform ( )

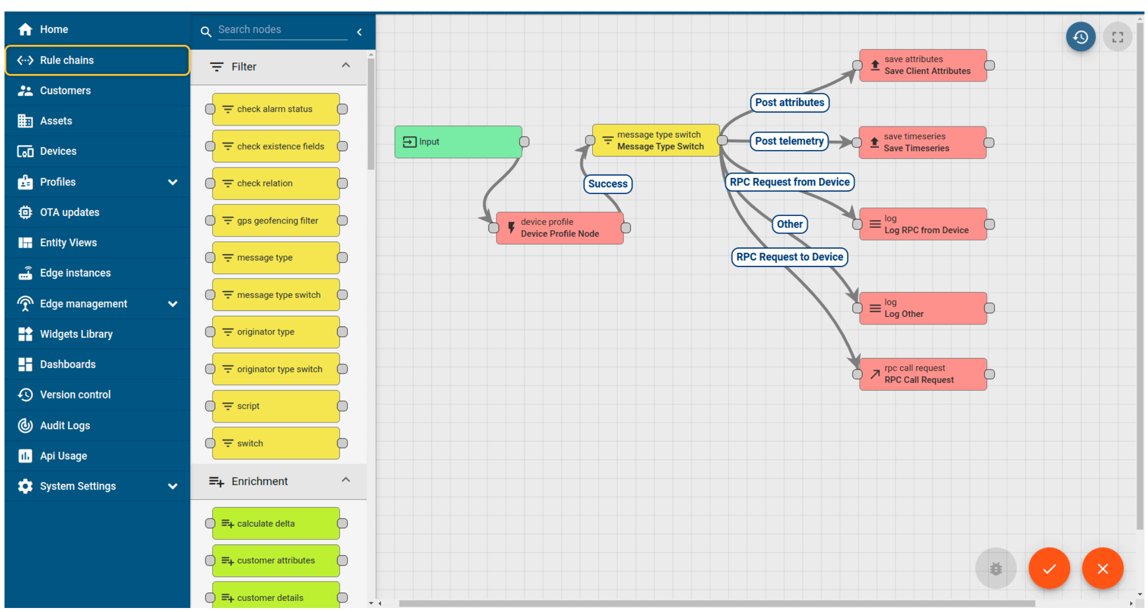
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable “insight” for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- 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





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.

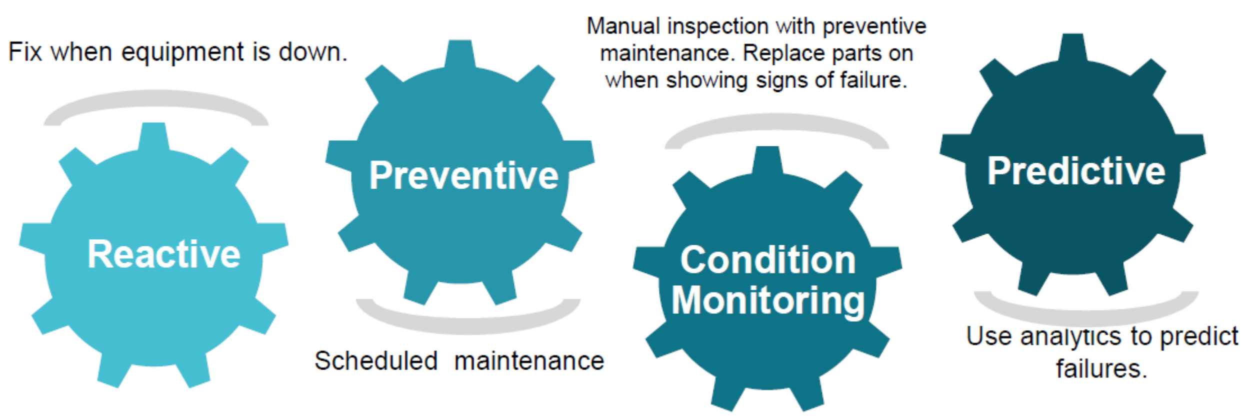
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1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

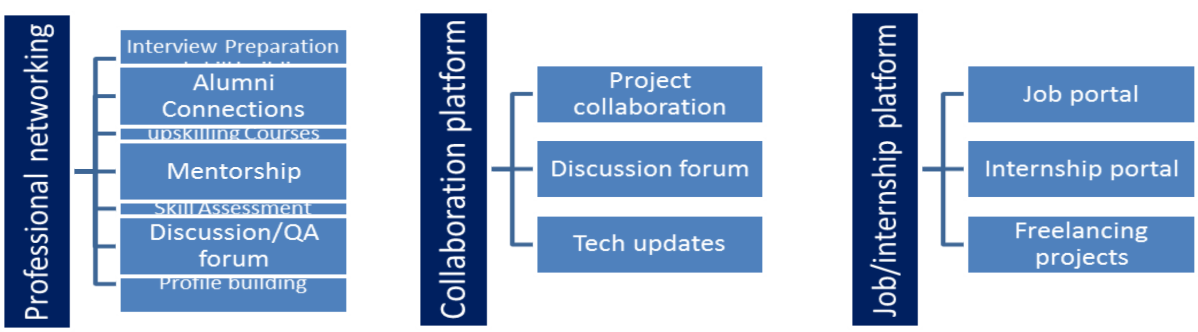
upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

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 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] (PDF) CROP YIELD PREDICTION BASED ON INDIAN AGRICULTURE USING MACHINE LEARNING (researchgate.net)

[2] Crop Production Prediction Using Machine Learning: An Indian Perspective | SpringerLink

[3] Crop yield prediction using machine learning: A systematic literature review - ScienceDirect

3 Problem Statement

In the assigned problem statement

Agriculture is a critical sector in India, providing livelihood to a significant portion of the population and contributing substantially to the national economy. Accurate predictions of crop production are essential for effective planning and decision-making, impacting everything from resource allocation and food security to market pricing and export strategies. Traditional methods of predicting crop yields often rely on historical data and heuristic approaches, which can be limited in accuracy and scope.

Challenges

1. **Climate Variability**: Unpredictable weather patterns, including changes in temperature, rainfall, and extreme weather events, significantly impact crop yields.
2. **Soil Diversity**: India's diverse soil types and their varying characteristics require detailed understanding and precise data for accurate predictions.
3. **Data Integration**: Integrating various data sources, such as historical yields, meteorological data, soil properties, and satellite imagery, is complex and requires sophisticated processing techniques.
4. **Resource Management**: Efficient use of water, fertilizers, and other inputs is critical, and inaccurate predictions can lead to suboptimal resource allocation.
5. **Technological Adoption**: There is a need for user-friendly tools and platforms that can deliver actionable insights to farmers and policymakers, especially in rural and resource-constrained areas.

Objective

To develop a robust, accurate, and scalable machine learning model that can predict crop production in India by integrating diverse data sources. The model aims to provide insights that help farmers optimize their practices, enable policymakers to make informed decisions, and ensure overall agricultural sustainability.

Specific Goals

1. **Data Collection and Integration**: Gather and harmonize historical crop production data, meteorological data, soil characteristics, socio-economic factors, and satellite imagery.
2. **Preprocessing and Feature Engineering**: Clean and prepare the data, selecting relevant features that influence crop yields.
3. **Model Development**: Develop and compare multiple machine learning models, including linear regression, decision trees, random forests, support vector machines, and neural networks, to identify the most accurate approach.
4. **Evaluation and Validation**: Use performance metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared to evaluate model accuracy and robustness.
5. **Prediction and Visualization**: Create a user-friendly platform for visualizing predictions through graphs, heatmaps, and an interactive dashboard.
6. **Recommendations**: Provide actionable insights and strategic recommendations based on model outputs to enhance agricultural productivity and resource management.

Expected Outcomes

1. **Accurate Predictions**: Improved accuracy in predicting crop yields, helping stakeholders make informed decisions.
2. **Enhanced Resource Management**: Optimized use of resources like water and fertilizers, leading to cost savings and sustainability.
3. **Policy Support**: Data-driven insights to support policy formulation and implementation in the agricultural sector.
4. **Farmer Empowerment**: Accessible tools and information for farmers to improve their practices and increase productivity.
5. **Future Adaptability**: A scalable model that can incorporate real-time data and adapt to future changes in climate and agricultural practices.

By addressing these challenges and achieving these goals, the project aims to significantly contribute to the efficiency and sustainability of Indian agriculture.

4 Existing and Proposed solution

4.1 Code submission (Github link)

https://github.com/Ajinka6885/upskillCampus/blob/06644430915f333dfb8138211e4c62826c8713ca/Crop%20prediction%20model%20.ipynb

4.2 Report submission (Github link) :

https://github.com/Ajinka6885/upskillCampus/blob/ea149e8f8e5b29a60035ebaf6a5ceed225238b70/Prediction%20of%20Agriculture%20Crop%20Production_Ajinkya_USC_UCT%20(1).docx

5 Performance Test

This is very important part and defines why this work is meant of Real industries, instead of being just academic project.

Here we need to first find the constraints.

How those constraints were taken care in your design?

What were test results around those constraints?

Constraints can be e.g. memory, MIPS (speed, operations per second), accuracy, durability, power consumption etc.

In case you could not test them, but still you should mention how identified constraints can impact your design, and what are recommendations to handle them.

6 My learnings

1. **Data Collection and Integration**:
   - Understanding the complexity of gathering diverse datasets, including historical crop production, meteorological data, soil characteristics, socio-economic factors, and satellite imagery.
   - Learning to handle large datasets and integrate them effectively to create a comprehensive database for model training.
2. **Data Preprocessing**:
   - Mastering techniques for data cleaning, normalization, and feature selection to ensure the quality and relevance of the data.
   - Dealing with missing values, outliers, and inconsistencies in the data.
3. **Machine Learning Algorithms**:
   - Gaining hands-on experience with various machine learning models, such as linear regression, decision trees, random forests, support vector machines (SVM), and neural networks.
   - Understanding the strengths and limitations of each model and learning to select the appropriate model based on the problem context.
4. **Model Training and Evaluation**:
   - Learning to split data into training and testing sets and using cross-validation to ensure model generalizability.
   - Using performance metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared to evaluate model accuracy and robustness.
5. **Prediction and Visualization**:
   - Developing skills in creating visualizations, such as graphs and heatmaps, to compare predicted vs. actual crop yields.
   - Building an interactive dashboard to allow users to explore predictions based on various parameters.
6. **Advanced Techniques**:
   - Exploring advanced techniques such as deep learning, ensemble methods, and explainable AI to enhance prediction accuracy and transparency.

Practical Learnings

1. **Interdisciplinary Knowledge**:
   - Appreciating the interdisciplinary nature of the project, which combines agriculture, meteorology, data science, and machine learning.
   - Understanding the real-world challenges faced by farmers and policymakers in the agricultural sector.

2. **Stakeholder Engagement**:
    - Learning to communicate effectively with stakeholders, including farmers, policymakers, and agricultural experts, to gather insights and validate model predictions.
    - Understanding the importance of making technology accessible and user-friendly for end-users, especially in rural and resource-constrained areas.
3. **Project Management**:
    - Gaining experience in managing a complex project, from data collection and preprocessing to model development, evaluation, and deployment.
    - Learning to set realistic goals, manage timelines, and coordinate with team members and collaborators.
4. **Ethical and Social Considerations**:
    - Understanding the ethical implications of using predictive models in agriculture, including the potential impact on farmers' livelihoods and the environment.
    - Considering data privacy and security issues when handling sensitive agricultural data.
5. **Policy and Economic Implications**:
    - Learning how accurate crop production predictions can influence agricultural policies, market prices, resource allocation, and overall economic stability.
    - Understanding the broader socio-economic context in which agricultural decisions are made.
6. **Future Directions**:
    - Identifying potential areas for future research and development, such as incorporating real-time data, expanding geographic and crop coverage, and developing advanced predictive models.
    - Recognizing the importance of continuous learning and adaptation to stay updated with technological advancements and changing agricultural practices.

This project has provided a comprehensive learning experience, combining technical skills in data science and machine learning with practical insights into agriculture and stakeholder engagement.

7 Future work scope

1. **Real-Time Data Integration**:
   - **Weather Data**: Incorporate real-time meteorological data to provide dynamic and timely predictions.
   - **Market Prices**: Include real-time market data to help farmers make better decisions about when to plant and harvest based on market conditions.
   - **Remote Sensing**: Use real-time satellite imagery to monitor crop health and growth stages continuously.
2. **Advanced Modeling Techniques**:
   - **Deep Learning**: Employ advanced deep learning techniques, such as Convolutional Neural Networks (CNNs) for analyzing satellite images and Recurrent Neural Networks (RNNs) for time-series forecasting.
   - **Ensemble Methods**: Combine multiple models to improve prediction accuracy and robustness.
   - **Explainable AI**: Develop models that not only predict but also provide explanations for their predictions to improve transparency and trust.
3. **Expanded Dataset**:
   - **Geographic Coverage**: Extend the model to cover more regions, including underrepresented and marginal areas.
   - **Crop Diversity**: Include a wider variety of crops, focusing on both major and minor crops to provide comprehensive coverage.
4. **User Interface and Accessibility**:
   - **Mobile Applications**: Develop mobile-friendly interfaces to make predictions accessible to farmers in remote areas.
   - **Language Support**: Provide multilingual support to cater to diverse linguistic groups across India.
5. **Collaboration and Partnerships**:
   - **Government Agencies**: Collaborate with governmental bodies for better data access and implementation of policy recommendations.
   - **Research Institutions**: Partner with academic and research institutions for continuous model improvement and innovation.
   - **Private Sector**: Engage with agri-tech companies to integrate predictive models with existing agricultural tools and platforms.
6. **Policy Impact**:
   - **Scenario Analysis**: Develop tools for policymakers to simulate different scenarios and their potential impact on crop production.
   - **Resource Allocation**: Use predictions to optimize the allocation of resources such as water, fertilizers, and subsidies.
7. **Educational Outreach**:
   - **Training Programs**: Conduct training sessions and workshops for farmers and agricultural officers to help them understand and utilize the prediction models effectively.
   - **Awareness Campaigns**: Increase awareness about the benefits of using predictive models in agriculture through outreach programs and demonstrations.

Scope

1. **Enhanced Agricultural Productivity**:
    - Improved prediction accuracy will enable better planning and decision-making, leading to increased crop yields and overall agricultural productivity.
2. **Sustainable Farming Practices**:
    - By optimizing resource use based on precise predictions, farmers can adopt more sustainable practices, reducing environmental impact and enhancing long-term soil health.
3. **Economic Benefits**:
    - Accurate crop predictions can help stabilize market prices, reduce waste, and increase profitability for farmers by aligning supply with market demand.
4. **Food Security**:
    - Reliable predictions contribute to food security by ensuring adequate production and reducing the risk of crop failures due to unforeseen climatic conditions.
5. **Climate Adaptation**:
    - The model can be used to develop strategies for climate change adaptation, helping farmers adjust their practices in response to changing weather patterns.
6. **Data-Driven Policy Making**:
    - Policymakers can use prediction data to formulate and implement more effective agricultural policies, subsidies, and support programs.
7. **Technological Advancement**:
    - The project can spur further research and innovation in the field of agricultural technology, encouraging the development of new tools and methods for precision farming.
8. **Global Application**:
    - While the initial focus is on India, the methodologies and models developed can be adapted and applied to other countries facing similar agricultural challenges, promoting global food security and sustainable farming practices.