**Industrial Internship Report on**

**Prediction of Agriculture Crop Production in India**

**Prepared by**

**[Student name]**

|  |
| --- |
| *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 3](#_Toc139702806)

[2 Introduction 4](#_Toc139702807)

[2.1 About UniConverge Technologies Pvt Ltd 4](#_Toc139702808)

[2.2 About upskill Campus 8](#_Toc139702809)

[2.3 Objective 9](#_Toc139702810)

[2.4 Reference 9](#_Toc139702811)

[2.5 Glossary 10](#_Toc139702812)

[3 Problem Statement 11](#_Toc139702813)

[4 Existing and Proposed solution 12](#_Toc139702814)

[5 Proposed Design/ Model 13](#_Toc139702815)

[5.1 High Level Diagram (if applicable) 13](#_Toc139702816)

[5.2 Low Level Diagram (if applicable) 13](#_Toc139702817)

[5.3 Interfaces (if applicable) 13](#_Toc139702818)

[6 Performance Test 14](#_Toc139702819)

[6.1 Test Plan/ Test Cases 14](#_Toc139702820)

[6.2 Test Procedure 14](#_Toc139702821)

[6.3 Performance Outcome 14](#_Toc139702822)

[7 My learnings 15](#_Toc139702823)

[8 Future work scope 16](#_Toc139702824)

# Preface

The experience of working with the USC/UCT team fully was really great, the team and the mentors were actually into this journey of six weeks internship Opportunity given by USC/UCT.

I had been explored to evolving technology Data Science and Machine Learning during this six-week internship opportunity given by USC/UCT.

The program was all well planned. The USC/UCT gave the real- world experience about the Machine learning and guided throughout the internship.



I had gone through all the e-resources provided by the organization and utilized every video sessions to upskill myself and serve the internship provided by USC/UCT.

Thanks to Nitin Tyagi Apurv sir who have helped me throughout this journey.

I would highly recommend this internship to all my juniors and peers to upskill themselves through real world experience provided by the USC/UCT.

# Introduction

## 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.



1. 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

 

1. **Smart Factory 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 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.

 

1.  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.

1. 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.



## 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

<https://www.upskillcampus.com/>

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



## 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.

## 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.

## Reference

[1] The reference for the data set was taken by the Kaggle website.

[2] The lecture and video sessions in the project plan provided by UCT was upto the mark to learn about the data science and machine learning

[3] The e-resources provided in the week 2 till week 5 were good enough to learn about the Python and data science concepts.

## Glossary

|  |  |
| --- | --- |
| Terms | Acronym |
| Dataset | DST |
| Preprocessing | PP |
| Training Model | TM |
| Testing Model | TESM |
| Accuracy | ACC |

# Problem Statement

Agriculture is a critical sector in India, contributing significantly to the country's economy and food security. However, predicting crop production is a complex task influenced by various factors such as weather conditions, soil quality, irrigation practices, pest infestations, and the use of fertilizers. Accurate predictions of crop production can help farmers, policymakers, and stakeholders make informed decisions, optimize resource allocation, and enhance food security.

The objective of this project is to develop a machine learning model that predicts the production of various crops in different regions of India. The model will utilize historical data on crop yields, weather patterns, soil conditions, and other relevant agricultural data. By analyzing this data, the model aims to provide accurate and timely predictions that can assist in strategic planning and decision-making in the agricultural sector.

# Existing and Proposed solution

**Traditional Statistical Models:**

Linear Regression: Often used for predicting crop yields based on historical data, weather conditions, and other factors. However, linear regression may not capture the complex relationships between variables in agricultural data.

Time Series Analysis: Techniques such as ARIMA (Auto Regressive Integrated Moving Average) models predict future crop yields based on past trends. These methods can struggle with handling non-linearities and multiple influencing factors.

**Enhanced Prediction System Using Random Forest Regressor and Decision Tree**

* Develop an accurate and robust machine learning model using Random Forest Regressor and Decision Tree to predict agricultural crop production in India.
* Integrate diverse data sources, including historical crop yields, weather data, soil information, and satellite imagery.
* Provide actionable insights and predictions to farmers, policymakers, and stakeholders.

**Key Components of the Proposed Solution:**

**Data Collection:**

* Historical Crop Data: Collect historical data on crop yields for various regions in India from government databases and agricultural research institutions.
* Weather Data: Integrate weather data (temperature, rainfall, humidity, etc.) from meteorological departments and weather stations.
* Soil Data: Gather soil quality data (pH, nutrient levels, moisture content) from agricultural surveys and soil testing labs.
* Remote Sensing Data: Utilize satellite imagery and indices such as NDVI for real-time monitoring of crop health.

**Data Preprocessing:**

* Clean and preprocess the collected data to handle missing values, outliers, and inconsistencies.
* Perform feature engineering to create meaningful input variables (e.g., lagged weather variables, soil quality indices).

**Model Development:**

* Exploratory Data Analysis (EDA): Conduct EDA to understand data distributions, correlations, and key influencing factors.
* Algorithm Implementation:
* Decision Tree Regressor: Implement a Decision Tree Regressor to create a simple model that can be easily interpreted. Decision Trees are good at handling categorical data and capturing non-linear relationships.
* Random Forest Regressor: Implement a Random Forest Regressor, an ensemble method that builds multiple decision trees and merges them to get a more accurate and stable prediction. It reduces overfitting and improves generalization.

**Model Training and Optimization:**

* Train the models using the preprocessed data.
* Optimize hyperparameters using techniques such as Grid Search or Random Search to improve model performance.

**Model Testing and Evaluation:**

* Evaluate the trained models on a separate test dataset using metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R²) to measure performance.
* Perform cross-validation to ensure the model's generalizability.

**Advantages of this Approach:**

* Improved Accuracy: Random Forest Regressor enhances prediction accuracy by reducing overfitting and capturing complex relationships.
* Robustness: Ensemble methods like Random Forest are more robust and generalizable compared to single models.
* Actionable Insights: Provides clear and actionable insights to farmers and policymakers to improve decision-making and resource allocation.
* Interpretability: Decision Trees offer interpretability, allowing stakeholders to understand the key factors influencing crop yields.

## Code submission (Github link)

<https://github.com/Anu123shree/upskillcampus/blob/main/crop_production_prediction.pdf>

## Report submission (Github link) :

# Proposed Design/ Model

Given more details about design flow of your solution. This is applicable for all domains. DS/ML Students can cover it after they have their algorithm implementation. There is always a start, intermediate stages and then final outcome.

## High Level Diagram (if applicable)

Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

## Low Level Diagram (if applicable)

## Interfaces (if applicable)

Update with Block Diagrams, Data flow, protocols, FLOW Charts, State Machines, Memory Buffer Management.

# 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.

## Test Plan/ Test Cases

## Test Procedure

## Performance Outcome

# My learnings

You should provide summary of your overall learning and how it would help you in your career growth.

# Future work scope

You can put some ideas that you could not work due to time limitation but can be taken in future.