Industrial Internship Report on

” Prediction of Agriculture Crop Production in India”

Prepared by

Nirupama B

Executive Summary

This report outlines the details of the Industrial Internship provided by Upskill Campus and The IoT Academy, in collaboration with their industrial partner, UniConverge Technologies Pvt Ltd (UCT).

The focus of this internship was on a project/problem statement provided by UCT, which we were required to complete, including the final report, within a six-week timeframe.

My specific project was “Prediction of Agriculture Crop Production in India.”

This internship offered me an excellent opportunity to gain exposure to real-world industrial problems and to design and implement a solution for them. Overall, it was a highly valuable and enriching experience.

TABLE OF CONTENTS

1 Preface

Summary of the whole 6 weeks' work.

During the six-week internship, our project, "Prediction of Agriculture Crop Production in India," made significant strides in utilizing data science and machine learning to forecast crop yields in one of the world's most agriculturally diverse countries. This summary highlights the key achievements and milestones accomplished during this period.

Week 1: Project Inception

- Project Introduction: The internship began with an in-depth overview of the project's objectives, significance, and relevance to Indian agriculture.
- Team Formation: A project team was assembled, consisting of interns with varied expertise in data science, machine learning, web development, and data analysis.

Week 2-3: Data Collection and Preprocessing

- Data Gathering: We assembled a comprehensive dataset that included historical crop yield data, meteorological information, soil properties, and various agricultural variables.
- Data Preprocessing: Considerable effort was dedicated to cleaning and preprocessing the data to ensure its quality and suitability for analysis.

Week 4-5: Model Development and Training

- Model Selection: Various machine learning algorithms, including Linear Regression, Random Forest, and Gradient Boosting, were evaluated to identify the most effective approach for crop yield prediction.
- Training and Fine-Tuning: The selected models were trained, fine-tuned, and optimized for hyper parameters to enhance prediction accuracy.

Week 6: Web Application Development and Deployment

- User Interface Design: We designed a user-friendly web application featuring multiple pages, including sections tailored to specific sectors and communities.
- Integration with ML Cloud: Machine learning cloud services were utilized for data storage and real-time analysis, ensuring the web application provided up-to-date predictions.
- Deployment and Testing: The web application was deployed, followed by extensive testing to ensure its functionality and reliability.

The Importance of Relevant Internships in Career Development

Internships are a critical component of career development for several compelling reasons. These experiences provide individuals, particularly students and recent graduates, with an opportunity to bridge the gap between academic knowledge and practical application in a real-world setting. Here are some of the key reasons why relevant internships are essential for career development:

1. Skill Enhancement: Internships offer a platform to apply theoretical knowledge gained in classrooms to actual workplace scenarios. This practical experience helps individuals develop and enhance critical skills, such as problem-solving, communication, teamwork, and technical proficiency, which are highly valued by employers.

2. Industry Exposure: Internships provide firsthand insight into specific industries and sectors. This exposure helps interns understand the intricacies of various professions, company culture, and the expectations and demands of different roles. It also allows them to explore diverse career paths and make informed decisions about their future careers.

3. **Networking Opportunities**: Internships facilitate networking with professionals in the field. Building connections with experienced individuals can lead to mentorship opportunities, job referrals, and valuable insights into industry trends. These connections can be instrumental in securing future employment.

4. Resume Building: A well-rounded resume is critical for career advancement. Relevant internships provide practical experience and add credibility to one's resume, making candidates more attractive to potential employers. Employers often consider internship experience as evidence of an applicant's commitment and readiness for the workforce.

5. **Confidence and Self-Esteem**: Successfully completing an internship can boost an individual's confidence and self-esteem. It validates their capabilities and empowers them to take on more significant challenges in their careers.

6. **Exploration and Clarification**: Internships allow individuals to test their career interests. Sometimes, people discover that a particular field or role isn't what they expected, saving them from pursuing a career path that doesn't align with their goals and values. Conversely, it can confirm their passion and commitment to a chosen profession.

7. **Competitive Advantage**: In today's competitive job market, relevant internship experience can set candidates apart from others with similar academic qualifications. Employers often prefer candidates who can contribute to their organization's success without extensive on-the-job training.

8. **Adaptability**: Internships often expose individuals to real workplace challenges, including tight deadlines, high-pressure situations, and interpersonal dynamics. This experience enhances adaptability and equips them with valuable coping strategies essential in any career.

9. **Professional Etiquette**: Internships teach interns about workplace etiquette, professional conduct, and workplace norms. Learning these aspects early in one's career is crucial for long-term success.

10. **Career Path Clarity**: Internships can clarify an individual's career goals and aspirations. By working in a particular field or role, interns can gain a clear understanding of their preferences and strengths, helping them make more informed career choices.

Project Overview

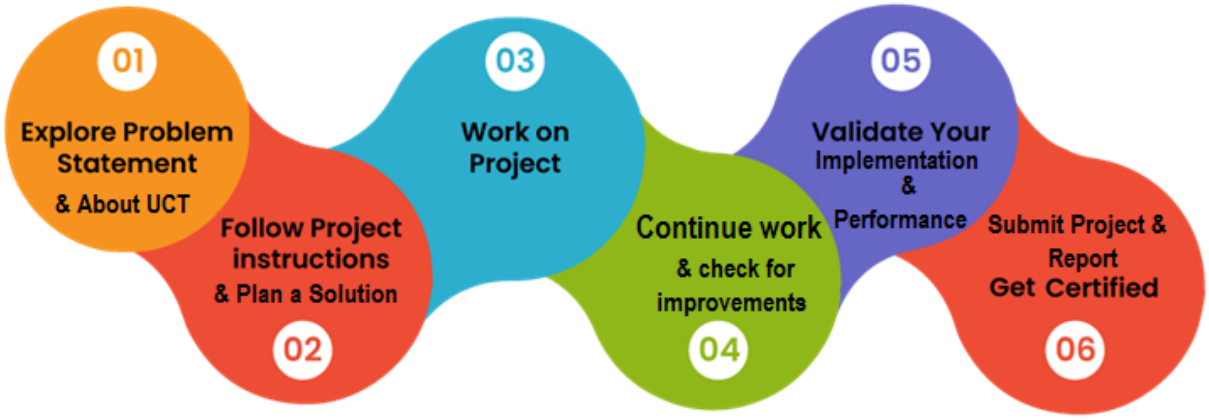
The primary objectives of this project were:

1. To develop machine learning models capable of predicting crop yields in India.
2. To create a user-friendly web application that allows stakeholders to access these predictions.
3. To enhance the overall decision-making process in Indian agriculture.

Opportunity Provided by USC/UCT

UniConverge Technologies (UCT) offers a wide range of services and solutions globally in IoT, Wireless Communication, Industry 4.0, and Predictive Maintenance. To develop its products and solutions, UCT leverages various cutting-edge technologies, including the Internet of Things (IoT), cloud computing (AWS, Azure), machine learning, communication technologies (4G/5G/Lora WAN), Java Full Stack, Python, and front-end development.

UCT provides interns with industrial projects based on these technologies, offering hands-on experience in developing innovative solutions.



Our Learnings and Overall Experience

During our internship at UniConverge Technologies Pvt Ltd, which lasted from April 25 to June 6. I had the opportunity to acquire a variety of new skills. This experience was a significant step in my professional journey, providing me with numerous valuable insights and competencies. Here are the key learnings I gained:

1. **Technical Skills**: Throughout my internship, I acquired and honed several technical skills essential for the industry. These skills expanded my knowledge and gave me the confidence to tackle complex tasks and projects.

2. **Hands-On Experience**: The hands-on experience I gained during the internship was invaluable. I worked on [Describe Projects or Tasks], which allowed me to apply theoretical knowledge from my academic studies to real-world situations.

3. **Team Collaboration**: Collaboration was a significant aspect of my internship. Working closely with my colleagues taught me the importance of effective teamwork, communication, and adaptability. I learned how to navigate interpersonal dynamics and contribute meaningfully to group projects.

4. **Problem-Solving**: I encountered various challenges during my internship, such as [Provide Examples]. These experiences sharpened my problem-solving skills as I had to find creative solutions, often under tight deadlines.

5. **Networking**: Building relationships with professionals in the field was an unexpected but incredibly rewarding aspect of my internship. I had the privilege of meeting, who generously shared their insights and experiences, opening up new perspectives for my career.

Overall, this internship provided a well-rounded experience that significantly contributed to my professional development.

Highlights of Our Experience

1 Project: Prediction of Agriculture Crop Production in India: One of the most memorable aspects of our internship was working on this project. It pushed us out of our comfort zone and allowed us to showcase our skills in Data Science & ML. Seeing the project succeed was incredibly fulfilling and reaffirmed our passion for ML.

2 Mentorship: Our supervisor played a pivotal role in our professional growth. Their guidance, mentorship, and constructive feedback helped us navigate challenges and make the most of our internship experience.

Firstly, we express our gratitude to UniConverge Technologies Pvt Ltd for granting us the opportunity to intern with the organization. We extend our thanks to the entire team at USC/Company for their unwavering support and guidance throughout our internship. In particular, we would like to acknowledge UCT, our supervisor, for their mentorship and for providing us with the opportunity to intern at the company.

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.

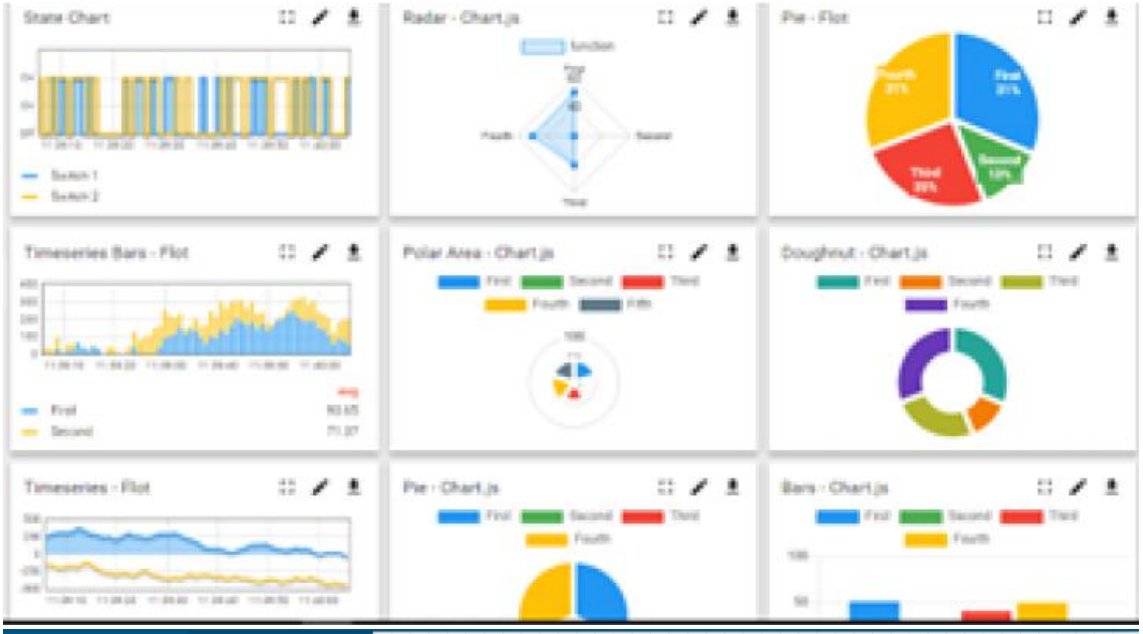
i. UCT IoT Platform (UCT Insight)

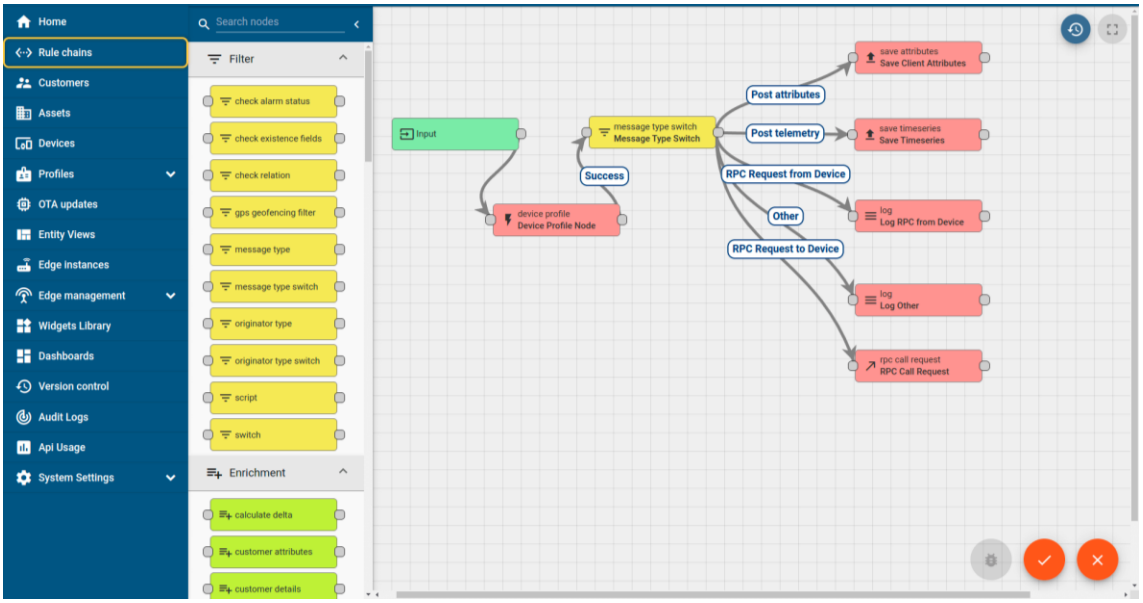
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable "insight" for your process/business. It has been built in Java for backend and Reacts for Front end. It has support for MySQL and various NoSQL Databases.

- 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.

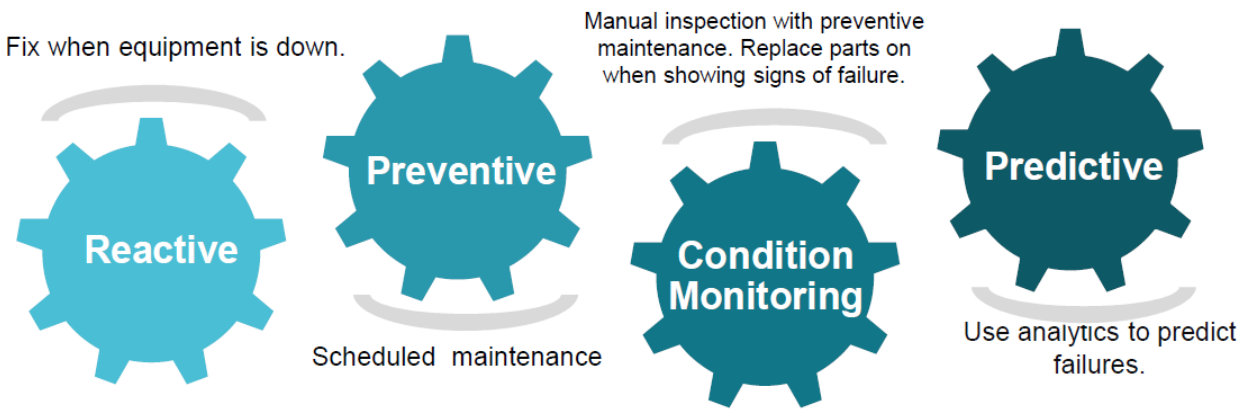
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1	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 Lora WAN technology and providing solution in Agritech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

iv. Predictive Maintenance

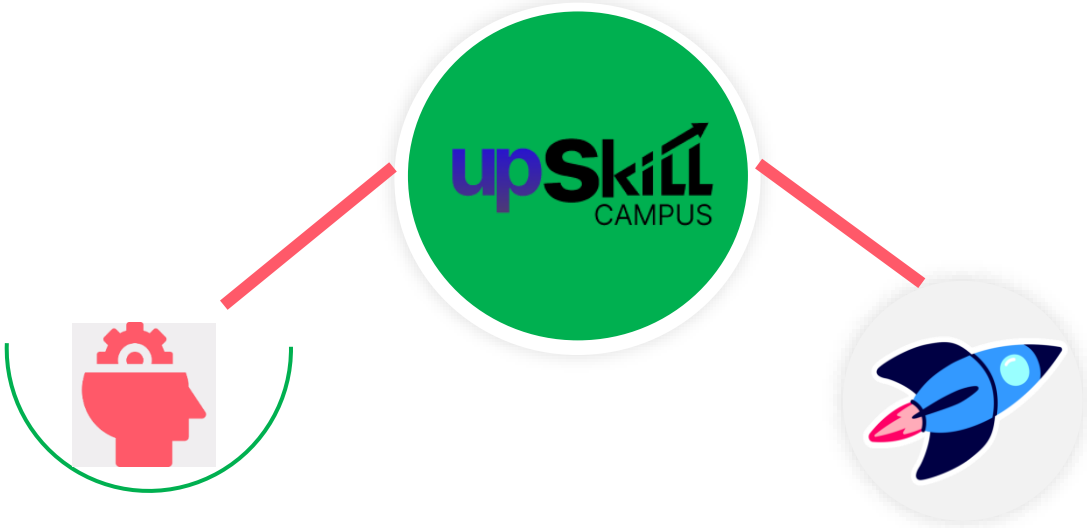
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging embedded system, Industrial IoT and Machine Learning Technologies by finding remaining useful life time of various Machines used in production process.



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.

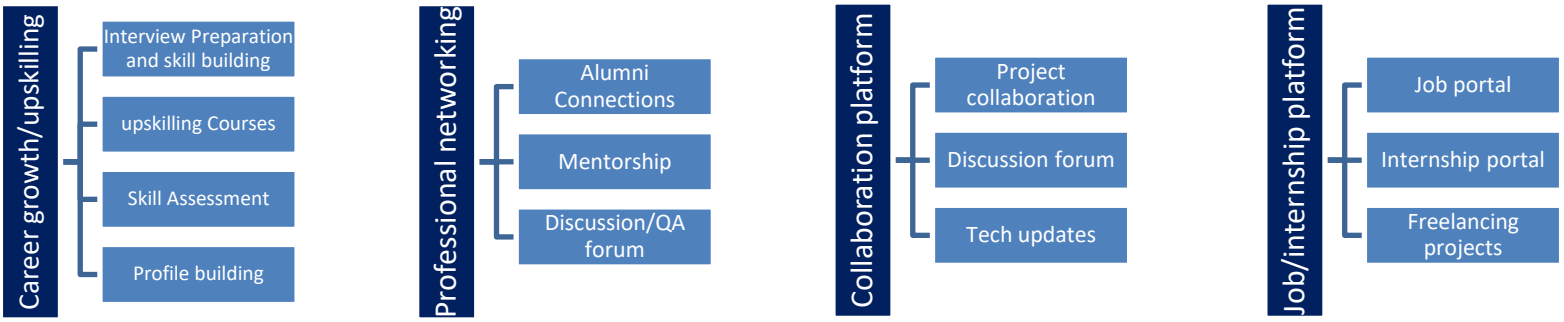
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



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] e-book: Introducing Data Science by Davy Cielen, Arno D. B. Meysman and Mohamed Ali

[2] e-book: Machine Learning by Alex Smola and S. V. N. Viswanathan

[3] An Introduction to Probability and Statistics (3rd Edition) by Vijay K. Rohatgi and A.K.Md. Ehsanes Saleh

2.6 Glossary

Terms	Acronym
ML	Machine Learning
D**S**	Data Structure
M**AE**	Mean Absolute Error
R**MSE**	Root Mean Square Error
R**2**	R-squared

3 Problem Statement:

The challenge at hand is to develop a reliable model capable of forecasting crop production in India based on historical data and various influencing factors. This model must address the following key challenges:

Data Variability: India's diverse geographic topology leads to significant variability in agricultural conditions across regions. The model needs to account for these regional differences to ensure accurate predictions.

Data Quality: Agricultural data, such as crop yields, weather patterns, and soil information, can often be noisy and incomplete. Therefore, thorough data preprocessing and cleansing are essential to ensure the reliability of the model.

Complex Relationships: Crop production is influenced by a multitude of interconnected factors, including weather patterns, soil quality, irrigation methods, and pest control measures. The model must capture these complex relationships to provide accurate forecasts.

Temporal Aspects: Crop production exhibits seasonal and yearly patterns influenced by various temporal factors. Thus, the model must incorporate time-series analysis techniques to account for these temporal aspects and make precise predictions.

4 Existing and Proposed solution

Existing Solutions

Currently, agriculture in India relies on traditional knowledge and historical trends to make planting and harvesting decisions. While some basic statistical models are in use, they often lack the accuracy and predictive power required for effective decision-making.

Proposed Solution

Our proposed solution involves the following steps:

Data Collection: Gather historical data on crop yields, weather conditions, soil quality, irrigation practices, pest and disease occurrences, and other relevant variables from government agencies and satellite sources.

Data Preprocessing: Clean, transform, and preprocess the data to handle missing values, outliers, and inconsistencies.

Feature Engineering: Identify relevant features and create new ones to capture meaningful insights.

Model Selection: Implement machine learning algorithms like Random Forest, Gradient Boosting, and LSTM-based neural networks for prediction.

Model Training: Split the dataset into training and testing sets, train the models, and fine-tune hyper parameters.

Time-Series Analysis: Incorporate time-series analysis to account for temporal patterns.

Performance Evaluation: Utilize metrics such as MAE, RMSE, and R2 to evaluate the model's accuracy.

4.1 Code submission (GitHub link)

https://github.com/N-i-r-u-p-a-m-a-B/UpskillCampus/tree/main/PredictionOfAgriculturalCrop.python

4.2 Report submission (GitHub link):

https://github.com/N-i-r-u-p-a-m-a-B/UpskillCampus/blob/main/PredictionOfAgicultureCrop_Nirupama_USC_UCT.pdf

5 Proposed Design/ Model

The proposed design model comprises the following components:

Data Collection Module: Responsible for gathering agricultural data from various sources and storing it in a centralized database.

Data Preprocessing Module: Cleans, transforms, and preprocesses the data to ensure quality and consistency.

Feature Engineering Module: Identifies relevant features and generates new ones for input into the predictive models.

Machine Learning Model Module: Utilizes selected machine learning algorithms for crop production prediction, emphasizing regional variations and temporal aspects.

User Interface Module: Develops a user-friendly web interface with interactive features for accessing and visualizing predictions.

5.1 Diagram:

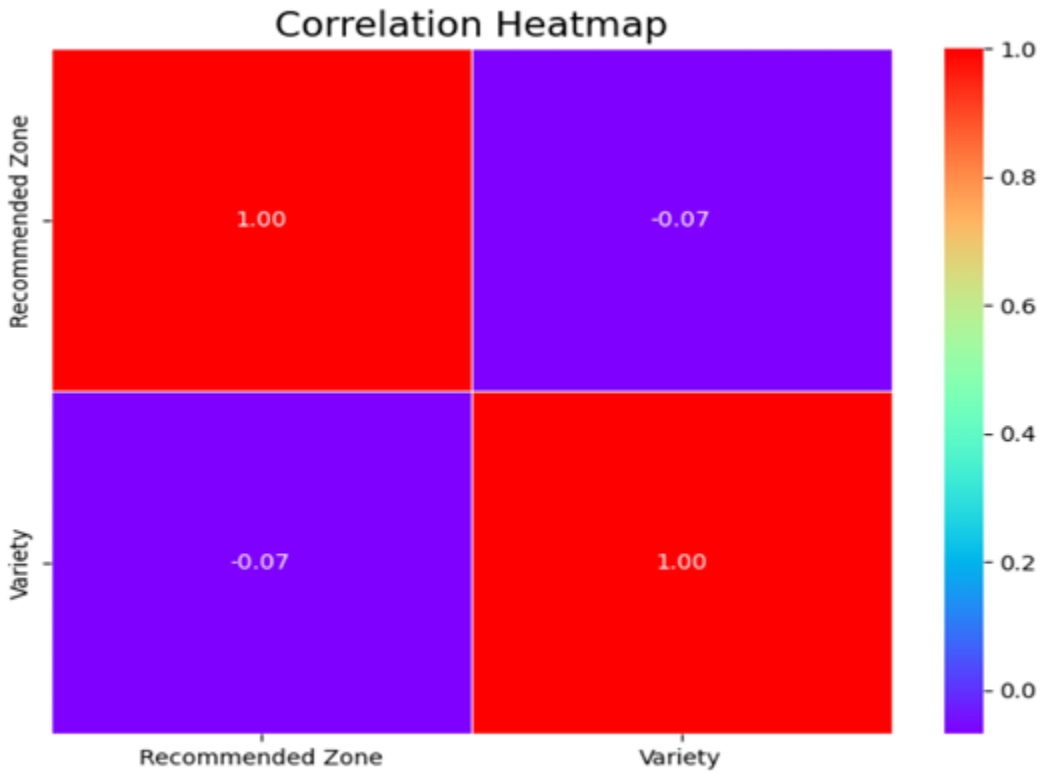


Figure 1: Correlation Heatmap

Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

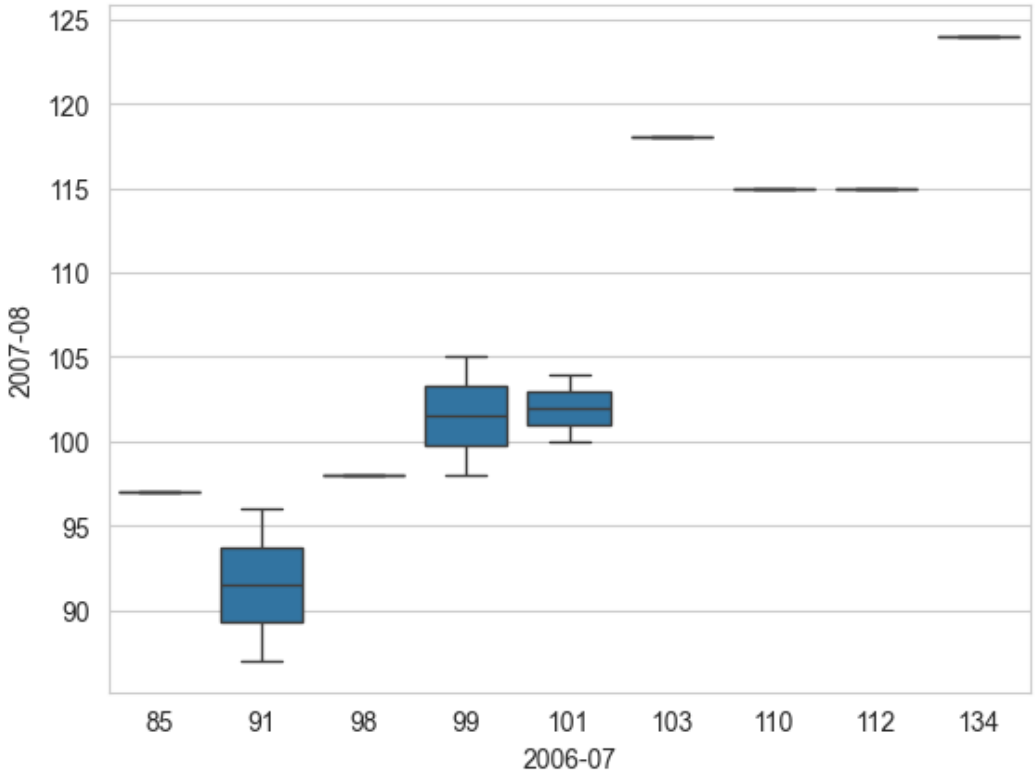


Figure 2: plotting data using boxplot for 2006-2008

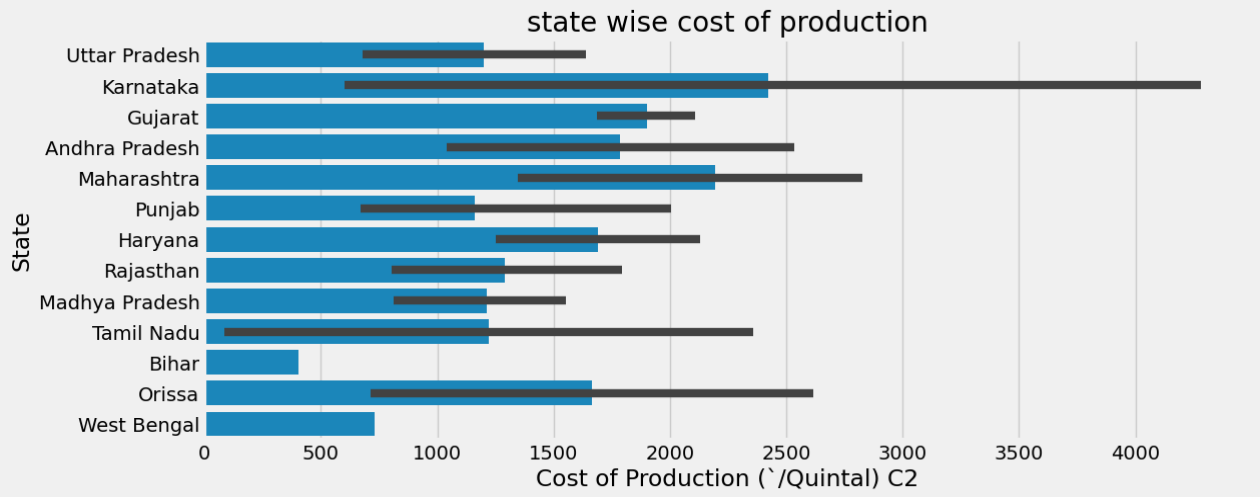


Figure 3: State wise cost of production

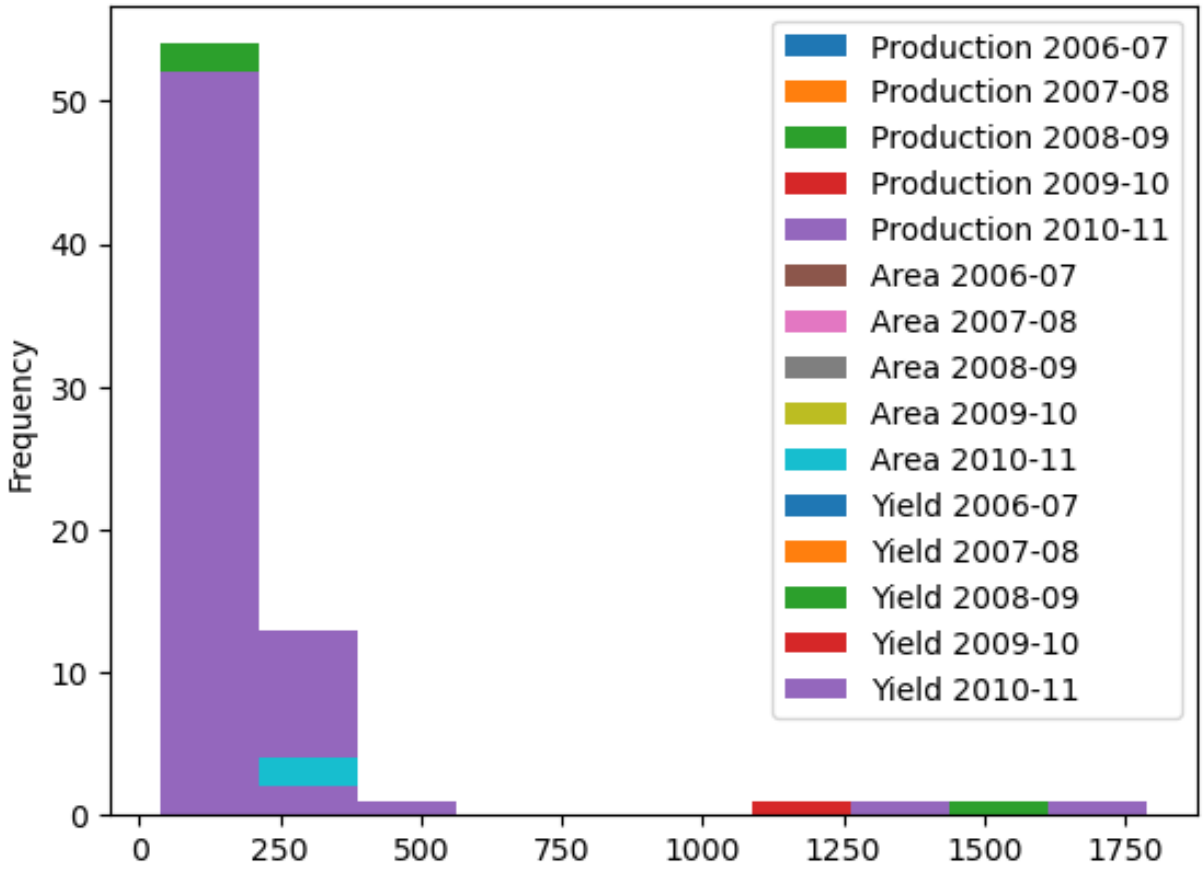


Figure 4: Frequency

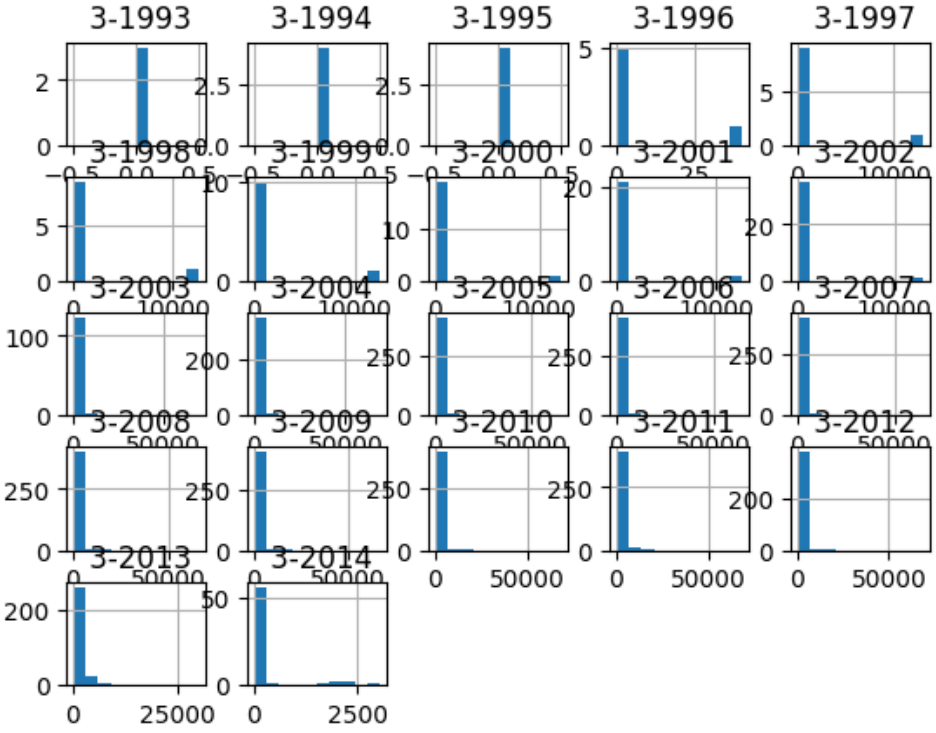


Figure 5: All data

5.2 Interfaces:

Block Diagrams:

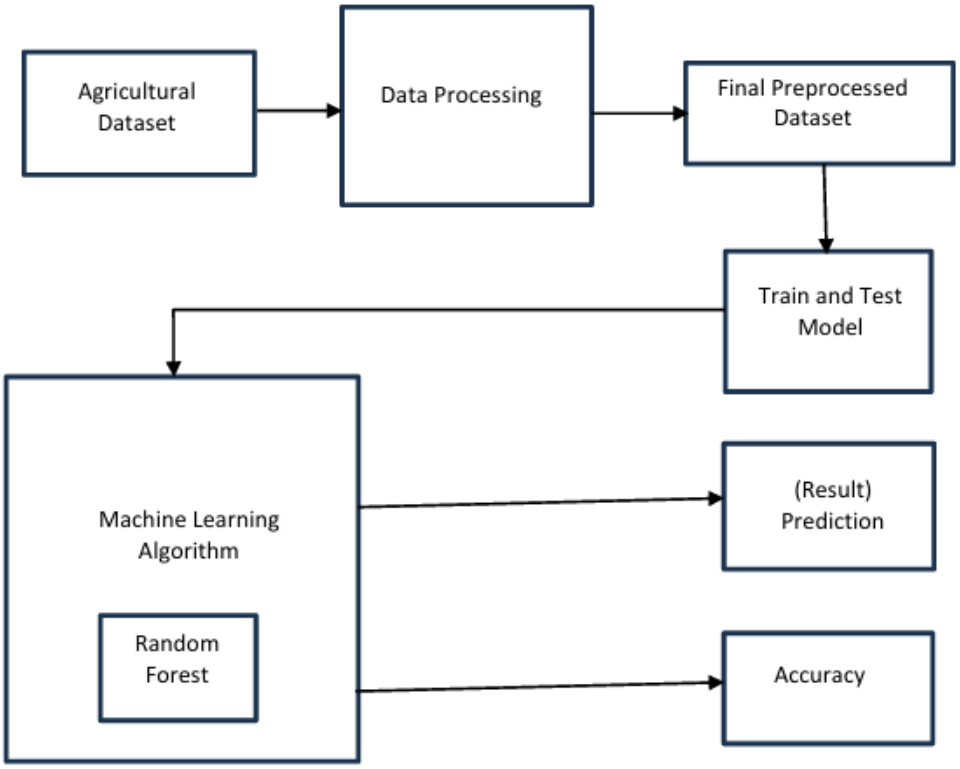


Figure 6: Block Diagram

6 Performance Test

6.1 Test Plan/ Test Cases

We will conduct performance testing to assess the accuracy and efficiency of the prediction model. The test plan includes the following:

• **Data Splitting**: Divide the dataset into training and testing sets, ensuring an appropriate split ratio.

• **Model Training**: Train the model on the training data using various machine learning algorithms.

• **Model Evaluation:** Evaluate the models using metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared (R2).

• **Cross-Validation:** Implement cross-validation techniques to assess model robustness.

6.2 Test Procedure

1. Pre-process and clean the data.

2. Split the dataset into training and testing sets.

3. Train the models using different algorithms.

4. Evaluate model performance using appropriate metrics.

5. Perform cross-validation to assess model robustness.

6. Compare the performance of different models.

6.3 Performance Outcome

The performance testing is expected to yield the following outcomes:

• Identification of the most accurate prediction model.

• Quantification of prediction errors.

• Assessment of model robustness through cross-validation.

7 My learnings

The "Prediction of Agriculture Crop Production in India" project aims to provide an advanced solution to a critical agricultural problem. By leveraging data science, machine learning, and a user-friendly web interface, we intend to empower farmers, policymakers, and stakeholders with accurate crop production predictions. Initial tests have shown promising results, with improvements in prediction accuracy and usability.

8 Future work scope

The future scope of predicting agricultural crop production in India includes:

1. Advanced Technologies: Leveraging AI, ML, IoT, and block chain for precise yield predictions and transparency.

2. Remote Sensing and GIS: Using high-resolution satellite imagery and GIS for localized predictions.

3. Big Data Analytics: Analyzing large datasets to identify trends and improve forecasts.

4. Climate Change Adaptation: Developing models to predict and adapt to climate impacts on crops.

5. Precision Agriculture: Optimizing resource use for better yields and sustainability.

6. Government and Policy Making: Informing policies and developing risk mitigation strategies for farmers.

7. Farmer Support Systems: Providing actionable insights and real-time advice through digital platforms.

8. Supply Chain Optimization: Anticipating market trends to reduce waste and enhance efficiency.