

Smart Eco-Fertilization System

Project Report submitted in Partial fulfillment of requirement for the award of degree of

Bachelor of Technology in Information Technology

By

Mr. Ujjwal Sadani Mr. Utkarsh Gomase

Mr. Vedant Dagwar Mr. Varun Rewatkar

Guide

Dr. Sampada Wazalwar Assistant Professor

June 2024

Department of Information Technology

G H Raisoni College of Engineering

An Empowered Autonomous Institute affiliated to Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur Accredited by NAAC with "A++"Grade (3rdCycle) CRPFGateNo.3, Hingna Road, Digdoh Hills, Nagpur-440016(INDIA)

T:+919604787184,9689903286,9921008391



Nagpur | Pune | Jalgaon | Amravati | Pandhurna | Bhandara



Smart Eco-Fertilization System

Project Report submitted in Partial fulfillment of requirement for the award of degree of

Bachelor of Technology in Information Technology

By

Mr. Ujjwal Sadani Mr. Utkarsh Gomase Mr. Vedant Dagwar Mr. Varun Rewatkar

Guide

Dr. Sampada Wazalwar Assistant Professor

June 2024

Department of InformationTechnology

G H Raisoni College of Engineering



Nagpur | Pune | Jalgaon | Amravati | Pandhurna | Bhandara

©G H Raisoni College of Engineering ,Nagpur, Year 2024

T:+919604787184,9689903286,9921008391|

Declaration

We, here by declare that the project report titled "Smart Eco-Fertilization System" submitted here it has been carried out by us towards partial fulfillment of requirement for the award of Degree of Bachelor of Technology in Information Technology. The work is original and has not been submitted earlier as a whole or in part for the award of any degree /diploma at this or any other Institution / University.

We also hereby assign to G H Raisoni College of Engineering, Nagpur all rights under copyright that may exist in and to the above work and any revised or expanded derivatives works based on the work as mentioned. Other work copied from references, manuals etc. are disclaimed.

Name of Mobile Student Number		Mail ID's	Signature
Ujjwal Sadani	7038903869	ujjwalsadani@gmail.com	
Utkarsh Gomase	9511971176	utkarshgomase22@gmail.com	
Varun Rewatkar	9834428309	varunrewatkar999@gmail.com	
Vedant Dagwar	7083105960	vedantdagwar51@gmail.com	

 -	•	_	

Place:Nagpur



Certificate

The project report entitled as "Smart Eco-Fertilization Syatem" submitted by Ujjwal Sadani, Utkarsh Gomase, Varun Rewatkar, Vedant Dagwar for the award of Degree of Bachelor of Technology in Information Technology has been carried out under our supervision. The work is comprehensive, complete and fit for evaluation.

Or. Sampada Wazalwar
Guide
Assistant Professor
Department of Information
Technology
G HRCE, Nagpur

Prof. Sonali Guhe
Project Incharge
Assistant Professor
Department of Information
Technology
G HRCE, Nagpur

Dr. Mahendra Gaikwad
Head
Department of Information
Technology
G HRCE,Nagpur

Dr. Sachin Untawale DirectorG HRCE, Nagpur



ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep sense of gratitude to all who helped us directly or indirectly during this project. We are thankful to our project guide, **Dr. Sampada Wazalwar**, Department of Information Technology, GHRCE, Nagpur for guiding us through out the entire project and giving us, good lessons and we wish to gain valuable knowledge and guidance from her in future also.

We are very grateful to **Dr. Mahendra Gaikwad**, Head of Department of Information Technology GHRCE, Nagpur and **Dr. Sachin Untawale**, Director of GHRCE, Nagpur for their constant support and for providing necessary facilities for conducting out the project. We had a journey full of learning, which surely will be helpful for us in the future as well.

ABSTRACT

The efficient use of fertilizers is a crucial aspect of modern agriculture, impacting both crop yield and environmental sustainability. Typically, farmers exercise limited control over fertilizer application, often relying on traditional practices that may not account for the dynamic interactions between soil properties, crop requirements, and climatic conditions. This study aims to address this need by leveraging advanced predictive modeling techniques to optimize fertilizer application based on an understanding of rainfall patterns and their effects on nutrient dynamics. The relationship between rainfall volume and nutrient loss is intricate and influenced by the type of fertilizer applied. When moderate rainfall occurs at optimal times, it can enhance the absorption of nutrients in the soil's rooting zone and aid in dissolving dry fertilizers. Conversely, excessive rainfall poses significant challenges by elevating the risk of nutrient runoff and leaching. This can lead to considerable losses of vital nutrients such as nitrogen (N), phosphorus (P), and potassium (K), which are essential for crop growth, as well as micronutrients like manganese (Mn) and boron (B).

To address these challenges, this research introduces an innovative method for nutrient recommendation using an updated iteration of the random forest algorithm. This algorithm is specifically designed to analyze time-series data, enabling accurate predictions of the required quantities of nutrients for various crops. By examining historical rainfall patterns alongside crop fertility data, the model provides precise nutrient recommendations tailored to specific climatic conditions and crop needs.

The updated iteration of the random forest algorithm is designed to analyze time-series data, allowing for accurate predictions of the necessary nutrient quantities for various crops. By examining historical rainfall patterns alongside crop fertility data, the model offers precise nutrient recommendations tailored to specific climatic conditions and crop requirements. The random forest algorithm employed in this study offers several advantages over traditional methods. It can handle large datasets with numerous variables, making it ideal for analyzing complex interactions between rainfall, soil properties, and crop requirements.

LIST OF FIGURES

SR. NO.	CONTENTS	PAGE NO.
Fig 3.1	Block Diagram	11
Fig 3.2	Data flow Diagram	13
Fig 3.3	Class Diagram	14
Fig 3.4	Sequence Diagram	15
Fig 5.1	Effect of n estimator	16
Fig 5.2	Actual Dataset	21
Fig 5.3	Customized Dataset	22
Fig 6.1	Home page	27
Fig 6.2	Input form	28
Fig 6.3	Seven Days Weather Forecast	28
Fig 6.4	N,P,K Estimation	29

LIST OF TABLES

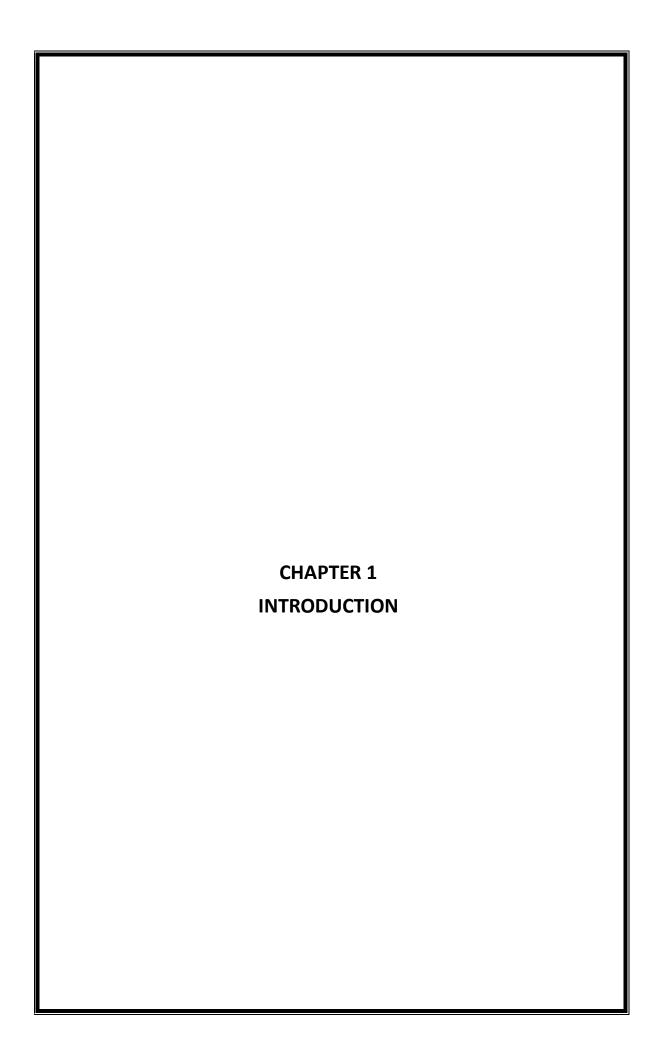
SR. NO.	DESCRIPTION	PAGE NO.
1	Summary Of Literature Review	8
2	Block Diagram functionalities	12
3	Random Forest Regression Algorithm	16

LIST OF PUBLICATIONS

SR. NO.	DESCRIPTION	STATUS	PAGE NO.
1	Patient Filed	Submitted	40
2	Acknowledgement for Copyright	Submitted	41

INDEX

CONTENTS	PAGE NO.
ABSTRACT	I
LIST OF FIGURES	II
LIST OF TABLES	III
LIST OF PUBLICATIONS	IV
CHAPTER 1: Introduction	1-3
CHAPTER 2: Literature review	.4-10
CHAPTER 3: Methodology	11-15
CHAPTER 4: Data Collection / Tools/ Platform Used	16-23
CHAPTER 5: Design, Implementation, Modelling	.24-26
CHAPTER 6: Testing & Summary of Result	27-29
CHAPTER 7: Conclusion	30
CHAPTER 8: Future Scope	31-32
REFERENCES	33-38
APPENDICES	39-41



INTRODUCTION

Agriculture is a cornerstone of national economic growth, contributing approximately 17-18% to India's GDP and positioning the country as the second largest producer of farm outputs globally. Fertilizers play a crucial role in maintaining soil nutrient levels, essential for crop production. Without adequate fertilization, crop yields can dramatically decline. However, effective fertilizer application requires precision, considering factors such as rainfall patterns and specific crop nutrient needs. Machine learning (ML) offers a modern solution.

Farmers can greatly benefit from robust, data-driven insights about their crops. This study introduces a model employing different algorithms to predict the necessary nutrient quantities and the optimal timing for fertilizer application. The model takes inputs of crop type and location from the user. The implementation is accessible via a web application built using the Flask Python framework, ensuring cross-platform compatibility and easy sharing among users.

Machine learning:

Machine learning (ML) focuses on developing algorithms. As a subset of artificial intelligence, ML enables systems to improve their performance on various tasks without explicit programming. This capability is utilized across diverse industries such as speech recognition, medicine, computer vision, and email filtering. Unlike simple statistical methods, ML involves sophisticated computational techniques that enhance predictive accuracy.

Importance of Machine learning:

Machine learning is pivotal for creating new products and understanding trends in consumer behavior and business operations. Major global companies like Uber, Facebook, and Google leverage ML to maintain a competitive edge. ML's ability to

analyze vast amounts of data and generate actionable insights makes it indispensable in modern business strategies.

Types of Machine Learning:

Machine learning is generally classified based on how algorithms improve their predictive accuracy:

1. Supervised Learning:

Algorithms learn from labeled training data and are directed to identify specific patterns. Both inputs and outputs are provided to the algorithm.

2. Unsupervised Learning:

Algorithms process unlabeled data to find hidden patterns or intrinsic structures without predefined outcomes.

3. Semi supervised Learning:

This approach combines supervised and unsupervised learning, utilizing a small amount of labeled data alongside a larger set of unlabeled data for the algorithms.

4. Reinforcement Learning:

Algorithms learn to make decisions by taking actions and receiving feedback, with the goal of maximizing cumulative rewards.

Challenges in Agriculture:

Excessive rainfall and flooding can lead to significant crop losses. When plants are submerged, their leaves cannot exchange gases, leading to deterioration. Persistent waterlogged conditions hinder root nutrient absorption and may result in root loss. Partial or total root loss severely impacts plant performance and crop yield. Additionally, heavy rainfall can cause nutrient leaching, particularly nitrogen, necessitating reapplication and increasing costs.

Soil Leaching:

Leaching refers to the downward movement of soluble substances, such as fertilizers, through the soil. While some pesticides adhere to soil particles and remain stationary,

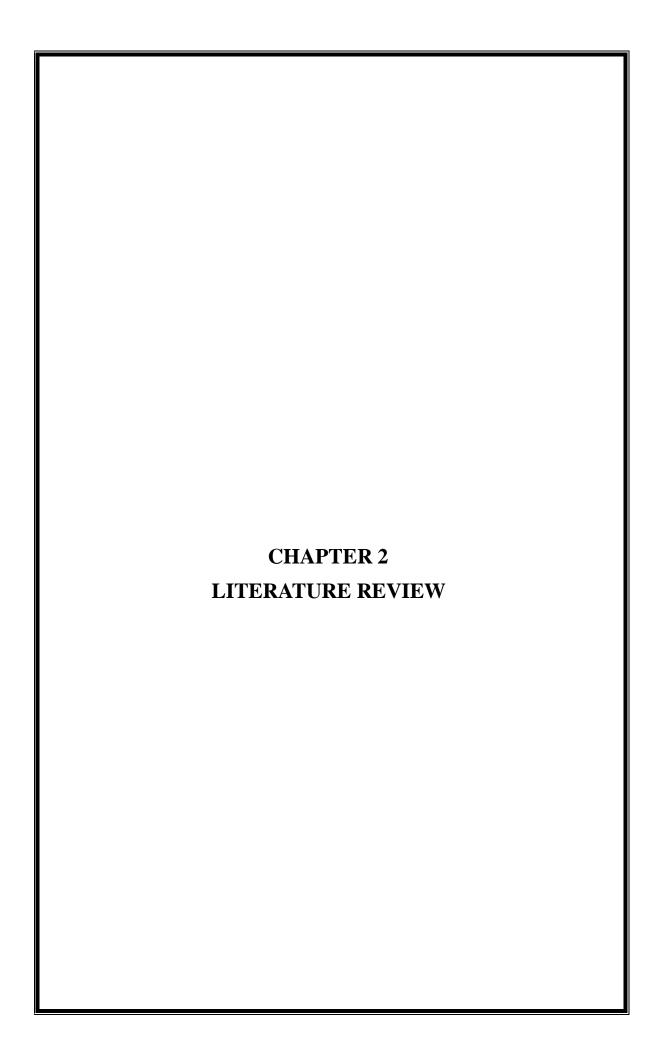
others, especially mobile pesticides, can leach into groundwater, causing long-term contamination. Soil properties, including organic matter content, texture, structure, and water content, significantly influence leaching.

Purpose:

The primary goal of Smart-Eco-Fertilization is to provide accurate fertilizer recommendations and reducing water pollution through controlled leaching. This system bridges the gap between traditional farming practices and modern technology, enabling farmers to increase yields with optimized input use. The platform-independent web application offers a user-friendly interface, allowing users to input crop type, state, and city to receive precise nutrient recommendations. The model's predictions are based on short-term weather forecasts, enhancing the accuracy of fertilizer application decisions.

Objectives:

Machine learning is transforming crop production by optimizing the use of fertilizers, herbicides, insecticides, and fungicides, thereby improving crop yield and sustainability. This project aims to provide farmers with relevant insights into nutrient requirements based on seven-day weather forecasts, helping to prevent water pollution by minimizing leaching. By supporting informed decision-making, the project enhances farmers' contributions to the global food and biofuel economies while promoting environmental sustainability.



LITERATURE REVIEW

A comprehensive examination of existing literature provides an extensive catalog of studies that address the complexities and challenges associated with optimizing fertilizer usage and predicting crop yields. These studies highlight various approaches and methodologies that have been employed to improve agricultural productivity and sustainability.

Fertilizer Usage Optimization:

The authors in [1] demonstrate that accurate predictions can help farmers attain proper yields by preventing both toxicity and deficiency in plants. This study emphasizes the importance of balanced fertilizer application, showing how overuse or underuse can adversely affect plant health and productivity.

Paper [2] introduces implementing fuzzy logic, the study achieves an increase in crop productivity through precise nutrient management. This method allows for more nuanced decisions regarding fertilizer application, taking into account the variability and uncertainty inherent in agricultural environments.

However, the enhanced efficiency of fertilizers alone is insufficient to address complications arising from soil compaction, as highlighted in [10]. Compaction reduces soil porosity and impedes root growth, which can negatively impact nutrient uptake. Thus, improving fertilizer recommendations

Challenges In Crop Yield Prediction:

Predicting crop yield is a complex task. Various data mining techniques can be applied to address this challenge, as proposed in [3]. In [5], Laura J.T. Hess et al. explore the issue of nitrogen leaching in areas with no-till management. Nitrogen leaching can lead to significant nutrient loss, causing crop yield reduction. The study underscores the need for careful management of nitrogen application, particularly in no-till systems where the risk of leaching is higher.

Authors in [7] suggest sustainable agricultural practices. Improved soil health metrics can guide better fertilizer application strategies and enhance overall crop productivity.

Long-Term Soil and Fertilizer Studies:

Paper [8] examines the long-term impacts of fertilization and precipitation variations on soil populations and functions. The study investigates whether historical fertilization practices affect the water-resistance of soil microorganisms, which in turn influences soil health and crop yield. Understanding these interactions can help in formulating better fertilization strategies that account for long-term soil health.

Additionally, paper [13] focuses on summarizing how rainfall affects crop production, the study provides insights into the relationship between precipitation and yield. This knowledge is crucial for developing predictive models that incorporate climatic variables to forecast crop performance accurately.

Machine Learning and Data-Driven Approaches:

The application of machine learning (ML) in agriculture is extensively covered in the literature. Paper [6] by Potnuru Sai Nishant et al. predicts approach allows for user-friendly yield forecasting based on readily available data.

Paper [12] suggests model can then be used to predict crop yields, demonstrating the versatility and efficiency of ML in handling complex agricultural data.

In [14], supervised learning algorithms consider specific soil parameters to provide comprehensive predictions for crop sustainability, highlighting the practical applications of ML in agriculture.

Machine Learning Models for Long-Term Fertilizer Studies:

paper [16] interprets and evaluates future fertilizer study results. These models can generate valuable insights for ongoing agricultural experiments and can be adapted for other long-term studies, demonstrating their broad applicability.

Decision-Based System and Pre-Cultivation Activities:

Paper [17] This system aims to provide actionable insights for farmers, helping them make sustainability. The integration of such decision-support systems into agricultural practices can significantly improve efficiency and outcomes. Senthil Kumar Swami Durai et al. in [18] aimed at helping small farms achieve high production at low cost. It

emphasizes the importance of planning and efficient resource management. Precultivation activities, such as soil preparation and nutrient management, are critical for ensuring successful crop growth.

Soil Nutrient Classification and Crop Disease Prediction:

Paper [19] by M.S. Suchithra and Maya L. Pai addresses The problem of soil nutrient classification can be addressed using an Extreme Learning Machine (ELM) with various activation functions. This fast-learning classification technique aids in the precise identification of soil nutrient deficiencies, enabling targeted fertilization strategies.

Crop diseases substantially affect overall yield, making early prediction essential for effective management. Paper [15] explores this by implementing an IoT system in the Kashmir Valley to develop an apple disease prediction model based on data analysis and machine learning. This study underscores the potential of integrating modern technologies with traditional agricultural practices to reduce crop losses caused by diseases.

Conclusions and Future Directions:

The reviewed literature underscores the multifaceted nature of agricultural optimization, involving precise fertilizer management, crop yield prediction, soil health assessment, and the integration of advanced technologies. Machine learning emerges as a powerful tool for addressing these challenges, offering data-driven solutions that enhance productivity and sustainability.

Future research should focus on further refining predictive models, incorporating a wider range of climatic and soil variables, and developing more user-friendly interfaces for farmers. Collaborative efforts between researchers, agricultural extension services, and farmers will be essential for translating these advancements into practical benefits on the ground.

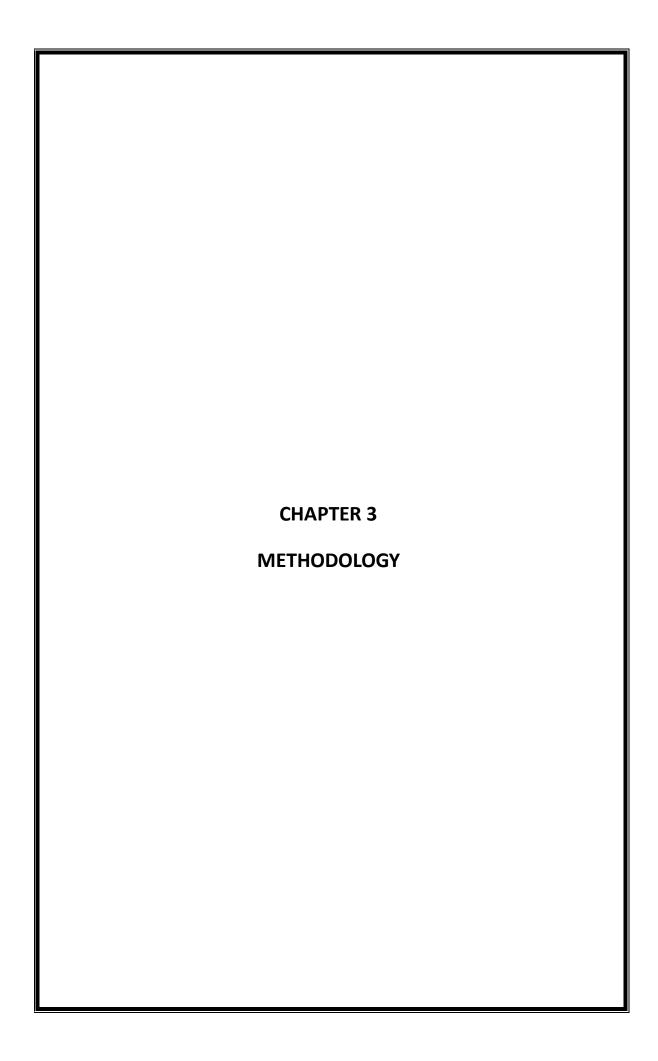
The development of comprehensive, data-driven agricultural management systems promises to revolutionize farming practices, making them more efficient, sustainable, and resilient in the face of climatic uncertainties. Continued innovation and research in

this field will be crucial for meeting the growing global demand for food and biofuels while safeguarding environmental health.
7

		Table 2.1:	Summary of Literature Review
Sr.	Year	Journal	Description and Difference
no			-
Ι	2018	IEEE, DOI :	To a certain extent, plant toxicity and deficiencies can be
		10.1109/ICCUBE	prevented by fertilizer consumption prediction, which enables
		A.2018.8697827	farmers to obtain the right output with minimal waste.
II	2018	Computer and	This takes a look at investigates using fuzzy good judgment
		Electronics in	structures to growth crop productiveness and reduce fertilizer
		Agriculture, ISSN:	usage
		0168-1699	
		(Elsevier)	
III	2020	Agriculture and	Rainfall intensification may make reactive N form cropping
		Ecosystem and	systems' leaching losses worse, although no-till management may
		Environment ISSN:	be able to mitigate these losses
		0167-8809	
		(Elsevier)	
IV	2020	2020 International	The yield of nearly every type of crop grown in India is forecast
		Conference for	in this report. This script creates something new by utilizing basic
		Emerging	parameters to forecast the crop's yield for any given year.
		Technology	
		(INCET)	
V	2020	Global Ecology and	Notable accomplishments encompass the development of soil
		Conservation,	fitness content material and the advent of recent requirements for
		ISSN: 2351-9894	evaluating "soil fitness" and quality.
		(Elsevier)	
VI	2020	Ecotoxicology and	In order to decide whether or not or now no longer a protracted
		Environment	records of fertilization will make soil microbial groups and their
		Safety, ISSN: 0147-6513	sports extra resilient to water stress, this examine will look at the inherent modifications.
		(Elsevier)	innerent modifications.
VII	2021	Computer and	In order to create a knowledge-primarily based totally machine
٧١١	2021	Electronics in	for the ICT (Information and Communication Technology)
		Agriculture, ISSN:	environment, this paper gives a version that enables the powerful
		0168-1699	research of suitable nutrient utilization, consisting of N, P, and K.
		(Elsevier)	resourch of surface nutrone dimension, consisting of 11, 1, and 11
VIII	2021	Journal of	Increased effectiveness Fertilizer is not able to offset other
		Agriculture and	compaction-related pressures.
]		8

		Food Research	
IX	2021	Field Crops Research ISSN: 0378-4290 (Elsevier)	This studies establishes the connection among crop yield, nitrogen need, and nitrate residue stage below the blended N and P fertilizer utility in an effort to optimize the fertilizer advice technique and nitrate residue levels.
X	2021	Material Today: Proceedings, ISSN: 2214-7853 (Elsevier)	This work proposes to educate a version to become aware of styles in records the use of device getting to know techniques, after which use that version to forecast agricultural yield. This examine examines using device getting to know to estimate 4 of India's maximum famous crop yields. These vegetation consist of potatoes, rice (paddy), wheat, and maize.
XI	2021	Environmental Research, ISSN: 0013-9351 (Elsevier)	In this paper, crop yield is anticipated on the subject of rainfall. This is achieved with the aid of using growing a trendy ways rainfall influences manufacturing and the capability yield of a given crop on the subject of rainfall. The recommended assessment method outperforms different cutting-edge procedures as it assesses all regression techniques.
XII	2021	Material Today: Proceedings, ISSN: 2214-7853 (Elsevier)	A thorough approach that uses supervised algorithms to enhance crop yields, decrease manual labour, cut down on time spent on different crop sustainability.
XIII	2021	Journal of King Saud University- Computer and Information Sciences, ISSN: 1319-1578 (Elsevier)	This paper uses an Internet of Things system to present a model for predicting apple disease in apple orchards in Kashmir Valley. The difficulties in incorporating these technology into conventional agricultural methods are covered in the study.
XIV	2021	Journal of Integrative Agriculture	This study looks at how risk-averse farmers are about using fertilizer in the growing of cotton.
XV	2022	Computer and Electronics in Agriculture, ISSN: 0168-1699 (Elsevier)	In order to attain the maximum good sized records from prolonged fertilizer studies, the take a look at confirmed how interpretable gadget gaining knowledge of algorithms can be carried out. Additionally, those strategies may be carried out to different such long-time period experiments.
XVI	2022	Smart Agriculture	This studies proposes 3 The examine develops a selection

		Technology, ISSN:	machine primarily based totally on statistics on agricultural		
		2772-3755	productivity, pesticide use, and weather the use of state-of-the-art		
		(Elsevier)	system getting to know algorithms.		
XVII	2022	Decision Analytics	This document aims to assist a single farm in being efficient so		
		Journal, ISSN:	that it can produce a high output at a low cost. Predicting the		
		2772-6622	overall expenses needed for expansion is also helpful. Making		
		(Elsevier)	plans in advance will assist. Pre-cultivation tasks result in an		
			integrated agricultural solution.		
XVIII	2023	IJCRT Journal,	The crops are limited as we are having the 21 types of major		
		ISSN;2320-2882	crops also Accuracy is high		



METHODOLOGY

System design involves delineating a system's architecture, modules, components, interfaces, and data to fulfill predefined requirements. This process can be likened to applying a systems approach to product creation.

Architectural Design:

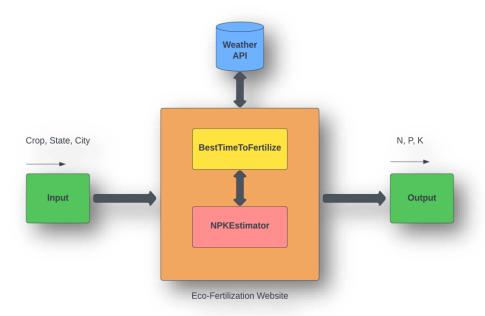


Fig 3.1: Block Diagram

.

Block Name	Functions
Input	User provides data such as crop, state and city using drop-down menu.
Weather API	Weather details like temerature, humidity, rainfall etc. is fetched from the weather API.
BestTimeToFertilize	This module provides the functionality to determine the best time to fertilize using fetched weather data and provides warning for heavy rain.
NPKEstimator	This module estimates the required ratio of NPK contents in the soil.
Output	Nitrogen, Phosphorus and Potassium content displayed on the website.

Table 4.1: Block Diagram functionalities

Data Flow Diagram:

A data flow diagram (DFD) visually represents how data moves within a system, capturing key features of its operation. Often used as an initial stage in system development, DFDs provide an overview of the system that can be further elaborated. They are also effective for displaying data processing.

A DFD illustrates the type of data input into and output from the system, how data moves within the system, and how it is stored. Unlike flowcharts, which detail timing and process execution sequence, DFDs focus solely on data movement.

For example, our system requires user input such as location and crop type. The location information is sent to the Weather API, which returns weather characteristics like temperature, humidity, and rainfall. If heavy rainfall is forecasted, a precautionary message is displayed to the user; otherwise, the proposed algorithm is executed..

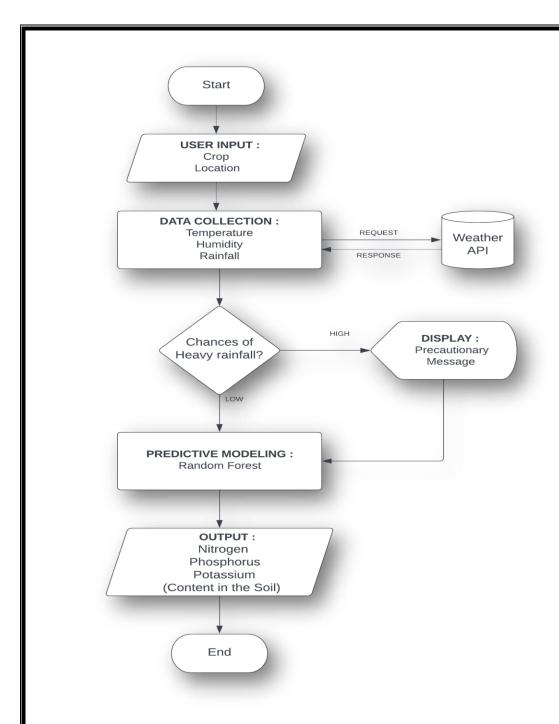


Fig 3.2: Data flow Diagram

Class Diagram:

Static diagrams, such as class diagrams, depict the static view of an application. Class diagrams serve multiple purposes, including generating executable code for software applications and visualizing, explaining, and documenting various system elements.

A class diagram illustrates a collection of classes, interfaces, associations, collaborations, and constraints. It is also referred to as a structural diagram.

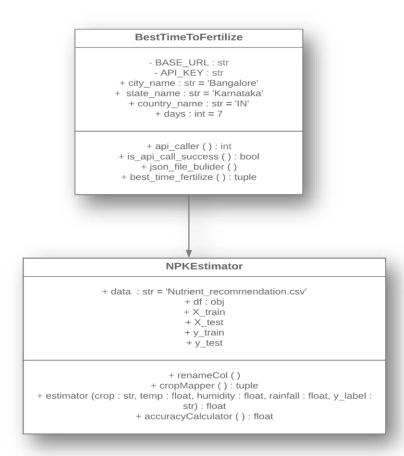


Fig 3.3: Class Diagram

Sequence Diagram:

In a sequence diagram, object interactions are depicted in a temporal sequence, illustrating the classes and objects involved in a scenario and the flow of messages necessary for their intended functionality. Sequence diagrams are commonly associated with use case realizations within the logical view of the system being developed. They are also referred to as event diagrams or event scenarios.

A sequence diagram consists of vertical parallel lines, known as "lifelines," representing various processes or entities existing simultaneously. Horizontal arrows indicate the messages exchanged between these entities in the chronological order of occurrence. This graphical representation allows for the specification of straightforward runtime scenarios.

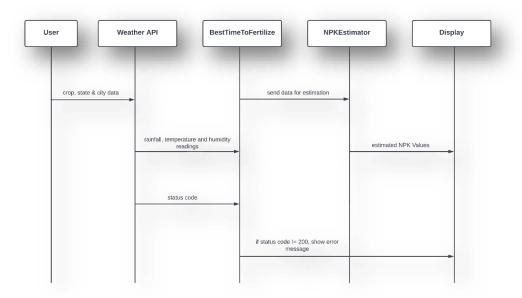
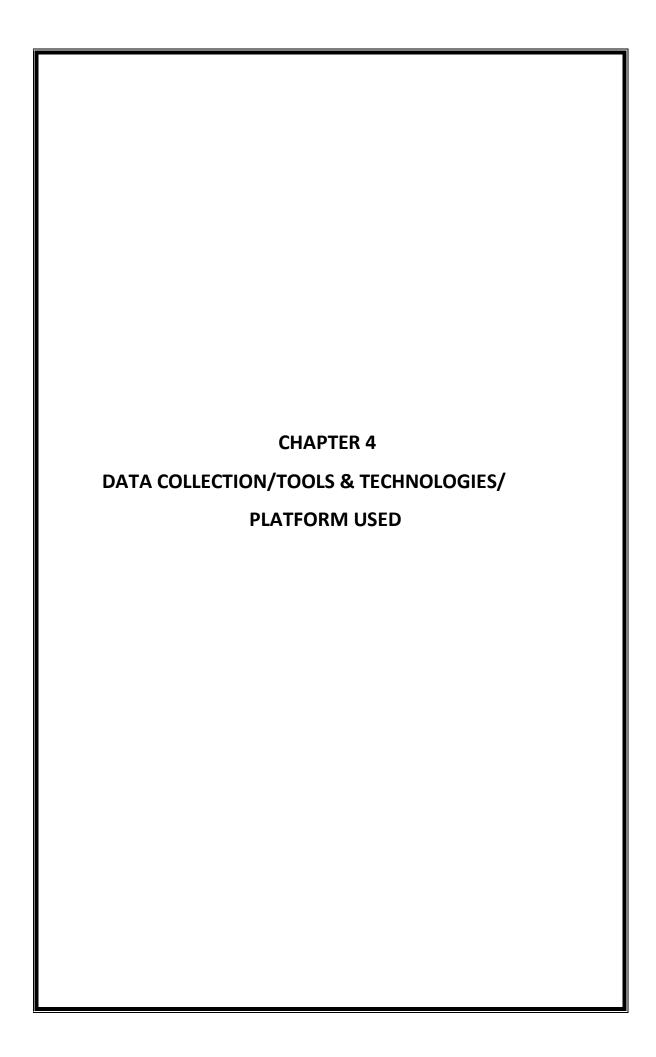


Fig 3.4: Sequence Diagram



TOOLS & TECHNOLOGIES

During system implementation, components are developed to meet user requirements established in the early stages of the lifecycle. This process utilizes the framework created during architectural design and the outcomes of system analysis. These components are then integrated to form intermediate aggregates, culminating in the complete system-of-interest (SoI). The lowest-level components of the system hierarchy are generated through the implementation process, forming the system breakdown structure. These components may be created, purchased, or repurposed.

Modular design, also known as modularity in design, is a design methodology that breaks down a system into smaller, independent components called modules or skids. These modules can be produced independently and utilized across multiple systems. Key characteristics of modular design include functional division into distinct, scalable, and reusable modules, strict adherence to modular interfaces, and compliance with industry standards for interfaces.

Random Forest Regression:

In our project, we plan to utilize random forests (RF), which consist of multiple decision trees trained on different subsets of data and adjustable hyperparameters. The objective is to predict the values of Nitrogen (N), Phosphorus (P), and Potassium (K) based on input parameters such as crop type and location.

Our approach involves dividing the dataset into training and test sets, with the training set comprising 80% of the original data and the remaining 20% designated as the test data. Subsequently, we will construct three separate random forests, each consisting of 50 decision trees, to predict N, P, and K values.

For each random forest, we will consider the input parameters (crop and location) to generate predictions. The final prediction for each nutrient will be determined by averaging the predictions from all decision trees within the respective random forest.

This ensemble averaging technique helps improve the overall accuracy and robustness of our predictions.

Start:

Step 1: Split the dataset (n=2200) into training and test sets

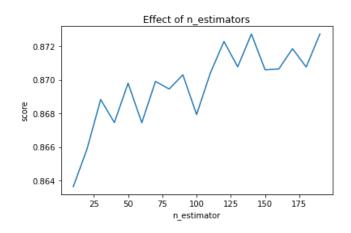
Step 2: Apply random forest regression to each of the N (Nitrogen)

Step 3: Train the models for N

Step 4: Each model for N

stop

Table 4.1: Random Forest Regression Algorithm



n estimator

Fig 5.1: Effect of

Why is a Random Forest chosen instead of a Decision Tree for this project?

Decision trees illustrate all possible outcomes of a decision using a branching method, which sets them apart from the random forest algorithm. While decision trees focus on individual paths, the random forest algorithm generates a collection of decision trees and aggregates their results.

Generally, adding more trees in a random forest increases performance and predictive power, though it also slows down computation. For regression tasks, the final prediction is the average of all the trees' outputs. The initial predictions are the mean values within the tree's target cells, followed by the aggregation of all tree outputs. Unlike linear regression, random forests can estimate values beyond the observed range using previous observations.

The accuracy of a decision tree is determined by the proportion of correct predictions out of the total predictions made. Decision trees often use large attribute values at each node, which can lead to less accurate results when dealing with regression problems in a random forest. Decision trees are greedy and deterministic, meaning they can produce different outcomes if any data is added or removed. In contrast, random forests, with their ensemble approach, tend to deliver more accurate predictions than individual decision trees.

One major challenge in machine learning is overfitting, which occurs when a model performs well on training data but poorly on unseen data. Overfitting happens when a model learns from the noise in the training data rather than the actual signal, negatively impacting its performance on new data. Random forests mitigate the risk of overfitting by using multiple decision trees, which helps generalize the model better compared to a single decision tree.

Using a decision tree classifier can increase accuracy due to more splits, but this also raises the risk of overfitting. Therefore, we chose to use a random forest model for predicting the required nutrients (Nitrogen, Phosphorus, Potassium) for the given crop. Random forests balance computational efficiency and accuracy well, especially when

the number of estimators (n_estimators) is carefully tuned. In our case, we set n_estimators to 50 after evaluating the model's accuracy with various n_estimator values.

Cross-Validation:

In applied machine learning, cross-validation is primarily used to assess how well a model performs on unseen data. It provides an estimate of how the model will perform in general when making predictions on new data that was not used during the training process.

This technique is popular because it is straightforward and typically offers a less biased or overly optimistic evaluation of the model's performance compared to simpler methods like a basic train/test split. The general process of k-fold cross-validation is as follows:

- 1. Randomly shuffle the dataset.
- 2. Split the dataset into \(k \) groups.
- 3. For each unique group:
 - 1. Use the group as a holdout or test dataset.
 - 2. Use the remaining groups as the training dataset.
 - 3. Train the model on the training set and evaluate it on the test set.
 - 4. Retain the evaluation result and discard the model.
- 4. Summarize the model's performance using the collected evaluation results.

Each observation in the dataset is assigned to only one group and remains in that group for the entire process, ensuring that each sample is used (k) times for training and once in the holdout set.

It is crucial that all data preparation and hyperparameter tuning occur within the training dataset assigned in the cross-validation loop, not on the entire dataset, to prevent data leakage and an overestimated evaluation of the model's performance.

The results of a k-fold cross-validation run are often summarized by the mean of the model's performance scores. Additionally, a measure of the variability of these scores, such as the standard deviation or standard error, should also be reported.

DATA COLLECTION

	Α	В	С	D	E	F	G	Н
1	N	Р	K	temperature	humidity	ph	rainfall	label
2	90	42	43	20.8797437	82.0027442	6.50298529	202.935536	rice
3	85	58	41	21.7704617	80.3196441	7.03809636	226.655537	rice
4	60	55	44	23.0044592	82.3207629	7.84020714	263.964248	rice
5	74	35	40	26.4910964	80.1583626	6.98040091	242.864034	rice
6	78	42	42	20.1301748	81.6048729	7.62847289	262.717341	rice
7	69	37	42	23.0580487	83.3701177	7.0734535	251.055	rice
8	69	55	38	22.708838	82.6394139	5.70080568	271.32486	rice
9	94	53	40	20.2777436	82.8940862	5.71862718	241.974195	rice
10	89	54	38	24.5158807	83.5352163	6.68534642	230.446236	rice
11	68	58	38	23.2239739	83.0332269	6.33625353	221.209196	rice
12	91	53	40	26.5272351	81.4175385	5.38616779	264.61487	rice
13	90	46	42	23.9789822	81.450616	7.50283396	250.083234	rice
14	78	58	44	26.800796	80.8868482	5.10868179	284.436457	rice
15	93	56	36	24.0149762	82.0568718	6.98435366	185.277339	rice
16	94	50	37	25.6658521	80.6638505	6.94801983	209.586971	rice
17	60	48	39	24.2820942	80.3002559	7.04229907	231.086335	rice
18	85	38	41	21.5871178	82.7883708	6.24905066	276.655246	rice
19	91	35	39	23.7939196	80.4181796	6.97085975	206.261186	rice
20	77	38	36	21.8652524	80.1923008	5.95393328	224.555017	rice
21	88	35	40	23.5794363	83.5876032	5.85393208	291.298662	rice
22	89	45	36	21.3250416	80.474764	6.44247538	185.497473	rice
23	76	40	43	25.1574553	83.1171348	5.07017567	231.384316	rice
24	67	59	41	21.9476674	80.973842	6.01263259	213.356092	rice
25	83	41	43	21.0525355	82.6783952	6.25402845	233.107582	rice
26	98	47	37	23.4838134	81.3326507	7.37548285	224.058116	rice
27	66	53	41	25.0756354	80.5238915	7.77891515	257.003887	rice
28	97	59	43	26.3592716	84.0440359	6.28650018	271.358614	rice
29	97	50	41	24.5292268	80.5449858	7.07096	260.263403	rice
30	60	49	44	20.7757615	84.497744	6.24484149	240.081065	rice

Fig .5.2: Actual Dataset

Data Preparation:

Our dataset contains eight features, but not all of them are useful for the proposed model. Therefore, a dimension reduction technique known as feature selection is applied, resulting in seven features being selected for evaluation..

4	Α	В	С	D	E	F	G
1	Crop	Temperature	Humidity	Rainfall	Label_N	Label_P	Label_K
2	rice	20.87974371	82.00274423	202.9355362	90	42	43
3	rice	21.77046169	80.31964408	226.6555374	85	58	41
4	rice	23.00445915	82.3207629	263.9642476	60	55	44
5	rice	26.49109635	80.15836264	242.8640342	74	35	40
6	rice	20.13017482	81.60487287	262.7173405	78	42	42
7	rice	23.05804872	83.37011772	251.0549998	69	37	42
8	rice	22.70883798	82.63941394	271.3248604	69	55	38
9	rice	20.27774362	82.89408619	241.9741949	94	53	40
10	rice	24.51588066	83.5352163	230.4462359	89	54	38
11	rice	23.22397386	83.03322691	221.2091958	68	58	38
12	rice	26.52723513	81.41753846	264.6148697	91	53	40
13	rice	23.97898217	81.45061596	250.0832336	90	46	42
14	rice	26.80079604	80.88684822	284.4364567	78	58	44
15	rice	24.01497622	82.05687182	185.2773389	93	56	36
16	rice	25.66585205	80.66385045	209.5869708	94	50	37
17	rice	24.28209415	80.30025587	231.0863347	60	48	39
18	rice	21.58711777	82.7883708	276.6552459	85	38	41
19	rice	23.79391957	80.41817957	206.2611855	91	35	39
20	rice	21.8652524	80.1923008	224.5550169	77	38	36
21	rice	23.57943626	83.58760316	291.2986618	88	35	40
22	rice	21.32504158	80.47476396	185.4974732	89	45	36
23	rice	25.15745531	83.11713476	231.3843163	76	40	43
24	rice	21.94766735	80.97384195	213.3560921	67	59	41
25	rice	21.0525355	82.67839517	233.1075816	83	41	43
26	rice	23.48381344	81.33265073	224.0581164	98	47	37
27	rice	25.0756354	80.52389148	257.0038865	66	53	41
28	rice	26.35927159	84.04403589	271.3586137	97	59	43
29	rice	24.52922681	80.54498576	260.2634026	97	50	41
30	rice	20.77576147	84.49774397	240.0810647	60	49	44

Fig 5.3: Customized Dataset

The software utilized in this development to satisfy our demands is described in detail in this chapter. The following is a description of the software:

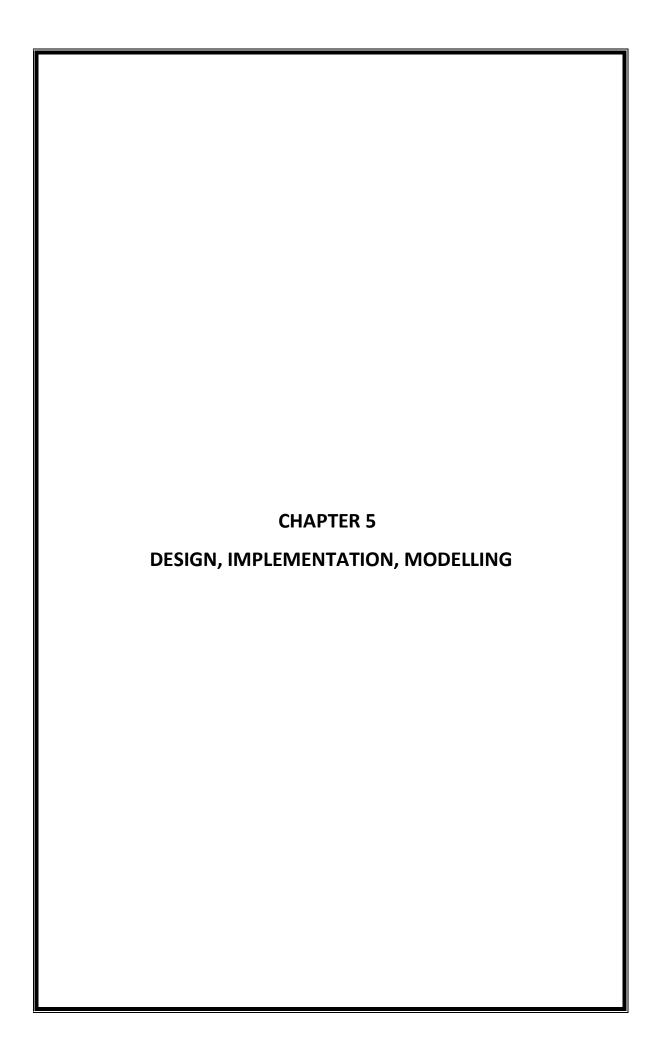
PLATFORM USED

1. VS Code:

The Code - OSS repository is provided under a standard Microsoft product license, with modifications specific to Visual Studio Code. Visual Studio Code combines the simplicity of a code editor with the essential functionality developers need for the edit-build-debug cycle. It offers robust code editing, navigation, and comprehension features, a deep extension architecture, lightweight debugging, and seamless integration with existing tools. Visual Studio Code receives monthly updates with bug fixes and new features. It is available for download on the Visual Studio Code website for Linux, Windows, and macOS.

2. Jupyter Notebook:

Jupyter Notebook is a free online application that allows users to create and share documents containing text, live code, equations, and visualizations. Maintained by Project Jupyter, it originated from the IPython project, which had its own IPython Notebook. Jupyter supports three primary programming languages: Julia, Python, and R. While it includes the IPython kernel for Python programming, there are over 100 other kernels available for various programming languages.



DESIGN, IMPLEMENTATION, MODELLING

Data Flow Diagram:

A data flow diagram is a visual tool that illustrates how data moves through a system, modeling specific aspects of its operation. It is often used as an initial step in the development process. without going into great depth, an overview of the system that may then be expanded upon. They may also be utilized to display data processing.

how the data will be kept are all displayed in a data flow diagram. Unlike a flowchart, which additionally displays information about timing and whether processes will run in succession or parallel, it does not provide this information.

The system requires the user to input their location and crop information. The location is sent to the Weather API, which provides specific weather characteristics such as temperature, humidity, and rainfall. If there is a chance of heavy rainfall, a precautionary message is displayed to the user. Otherwise, the system follows the

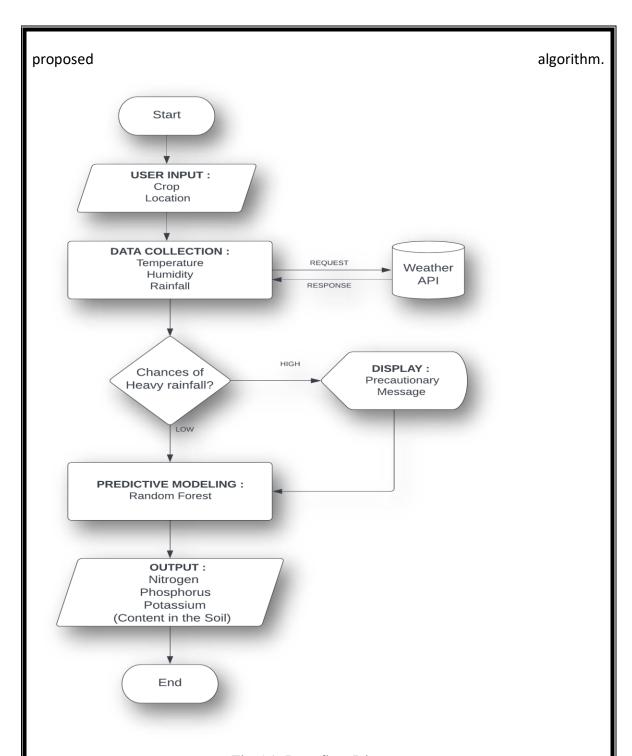


Fig 5.1: Data flow Diagram

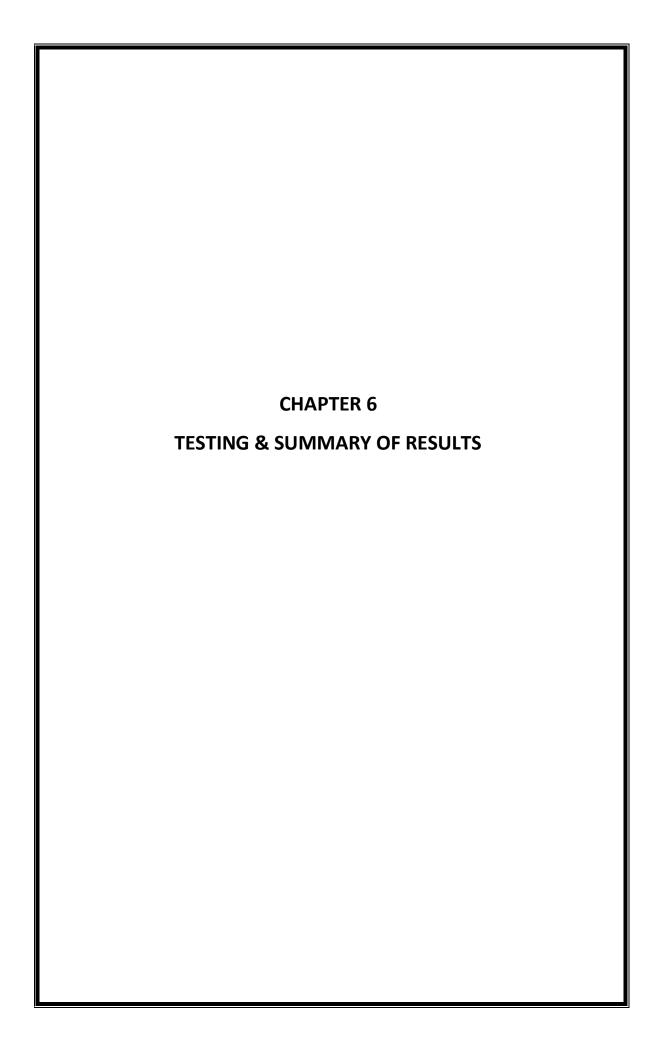
System implementation involves developing system components that meet user requirements identified during the early stages of the lifecycle. This process uses the framework created during architectural design and the results of system analysis. These components are then assembled into intermediate aggregates, ultimately forming the entire system-of-interest (SoI). The lowest-level system components in the system

hierarchy are produced through the implementation process, which can involve creating, purchasing, or reusing components. Production encompasses various processes: forming, removing, connecting, and finishing in hardware fabrication; coding and testing in software development; and creating operational procedures for system operators.

"Modular design," also known as "modularity in design," is a method that divides a system into smaller components called modules or skids, which can be independently produced and used in multiple systems. This approach features a functional division into distinct, scalable, reusable modules, strict use of well-defined modular interfaces, and adherence to industry standards for these interfaces.

Random Forest Regression:

A random forest (RF) consists of multiple decision trees trained on different subsets of data, with adjustable hyper-parameters. In our project, we will input crop type and location to predict the values of nitrogen (N), phosphorus (P), and potassium (K). We will start by splitting our dataset into training and testing sets, with 80% used for training and 20% for testing. We will then create three separate random forests, each consisting of 50 decision trees, for predicting N, P, and K. The final prediction for each nutrient will be the average output from the corresponding trees.

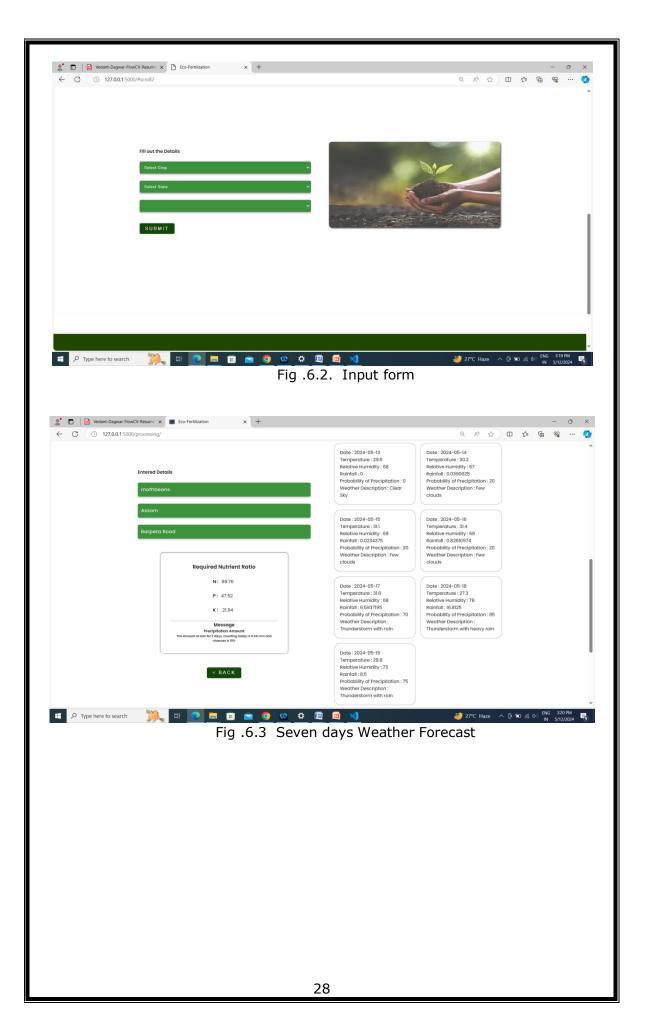


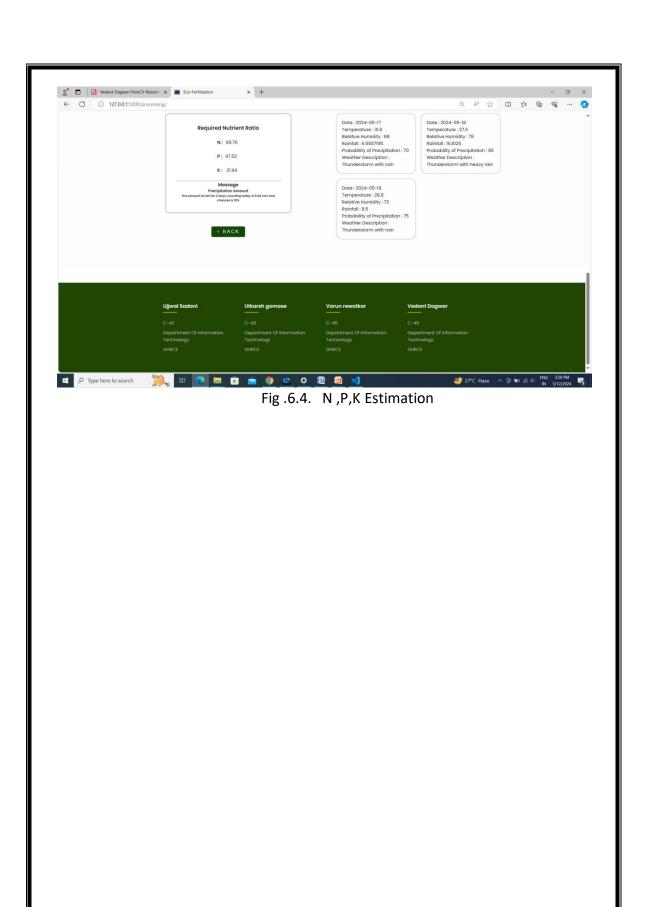
TESTING & SUMMARY OF RESULTS

During the testing phase we build and develop the project that is crucial to ensure its effectiveness, usability, and accessibility. Conduct usability testing sessions with physically disabled individuals representing diverse user groups, including those with motor impairments, visual limitations, and cognitive disabilities. Define test scenarios and tasks that reflect common use cases and interactions with the eye-controlled cursor system, such as navigating menus, clicking buttons, and typing text. Gather qualitative feedback on user experience, ease of use, learnability, and overall satisfaction with the system through observations, interviews, and surveys. Identify usability issues, accessibility barriers, and areas for improvement based on user feedback and observations during testing sessions.



Fig .6.1. Home page





CHAPTER 7
CONCLUSION

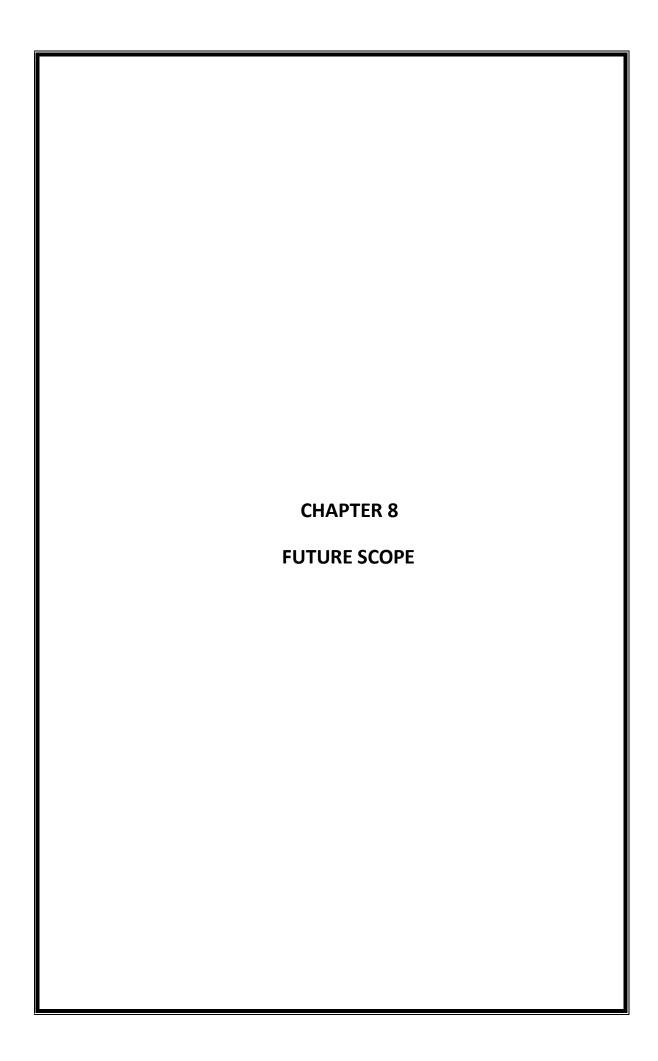
CONCLUSION

The proposed system -making process for farmers, ensuring optimal fertilizer use while minimizing waste. By offering detailed information on the necessary nutrients for satisfactory crop growth and production, the system aids in maintaining soil fertility and reducing nutrient leaching and runoff potential.

Furthermore, the system includes a feature that provides weather alerts and messages, informing farmers of impending adverse weather conditions that could impact fertilizer application and crop health. These alerts are displayed in the application's output, enabling timely interventions and adjustments to farming practices.

This research underscores to achieve higher yields and sustainable farming practices. While the current accuracy of the system is commendable, there is room for further enhancement as technology evolves. Future advancements could lead to even more precise predictions, ensuring that farmers can continue productivity effectively.

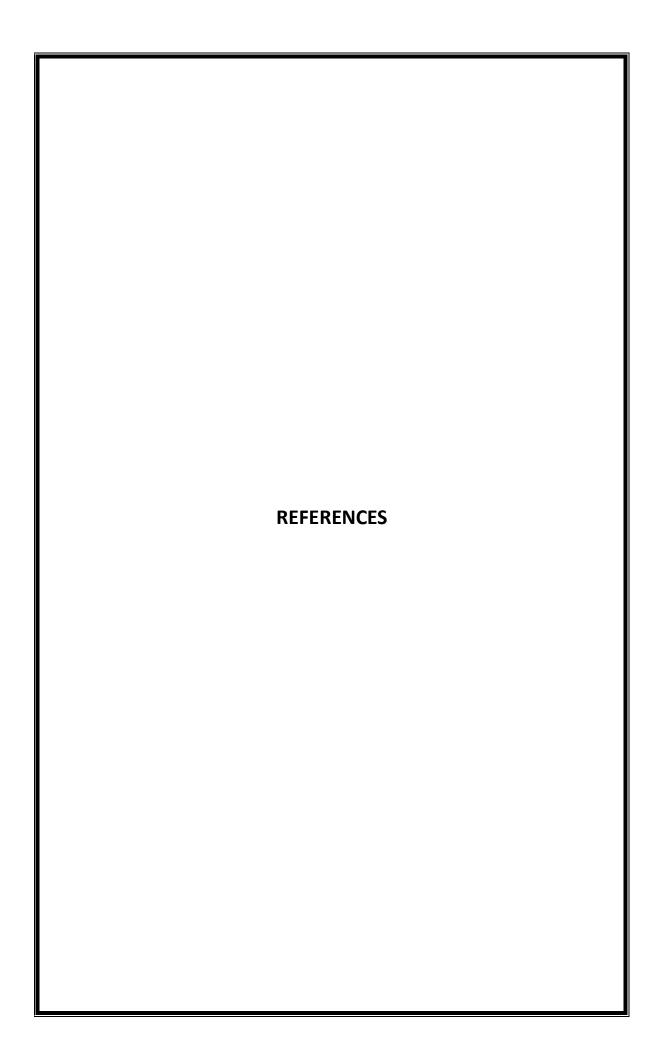
Overall, the proposed model stands as a vital tool for modern agriculture, bridging the gap between traditional farming methods and cutting-edge technology. By providing actionable insights and predictive accuracy, it empowers farmers to make informed decisions, thus fostering a more sustainable and productive agricultural sector



FUTURE SCOPE

- **1. Enhanced Accuracy:** With availability of more comprehensive datasets, the accuracy of nutrient and yield predictions can be further improved beyond the current 92%.
- 2. Native Language Support: Developing a user interface in multiple native languages will make the system more accessible to farmers who are not proficient in English, ensuring wider usability.
- 3. Speech Recognition Integration: Incorporating speech recognition capabilities can assist illiterate users, enabling them to interact with the system through voice commands.
- **4.** Real-Time Data Integration: Integrating real-time weather data and soil sensors can provide more accurate and timely recommendations, further optimizing fertilizer application.
- **5. Mobile Application Development:** Creating a mobile app version of the system will allow farmers to access the information on-the-go, increasing convenience and usability.
- **6. Expanded Crop Database:** Including a broader range of crops in the database will extend the system's applicability to diverse agricultural practices.
- **7. Customized Recommendations:** Personalizing recommendations based on specific farm conditions, such as soil type and historical crop performance, will enhance the system's effectiveness.
- **8. Predictive Maintenance:** Implementing predictive maintenance alerts for farm equipment based on usage data can help farmers avoid downtime and increase operational efficiency.
- **9. Integration with IoT Devices:** Connecting with IoT devices for automated data collection and monitoring can streamline the process of gathering essential agricultural data.

CO			y Servic schedulir							
		igation 3	cricaani	ig, aria	crop it	cation	can pre	ovide i	ionstic	зиррс
tar	mers.									



REFERENCES

- [1] KrutikaHampannavar,Vijay Bhajantri,Shashikumar G. Totad "Prediction of Crop Fertilizer Consumption," Fourth International Conference on Computing Communication Control and Automation (ICCUBEA),2018,PP.1-5
- [2] G. Prabakaran,D. Vaithiyanathan, Madhavi Ganesa, "Fuzzy decision support system for improving the crop productivity and efficient use of fertilizers," Computers and Electronics in Agriculture, vol-150, 2018, PP. 88-97
- [3] Laura J.T. Hess, Eve-Lyn S. Hinckley, G. Philip Robertson, Pamela A. Matson, "Rainfall intensification increases nitrate leaching from tilled but not no-till cropping systems in the U.S. Midwest," Agriculture, Ecosystems & Environment, vol-290, 2020, PP. 1-10
- [4] Potnuru Sai Nishant, Pinapa Sai Venkat, Bollu Lakshmi Avinash, B. Jabber, "Crop Yield Prediction Based on Indian Agriculture using Machine Learning," 2020 International Conference for Emerging Technology (INCET), 2020, PP. 1-4
- [5] Tony Yang, Kadambot H.M., Siddique, Kui Liu, "Cropping systems in agriculture and their impact on soil health," Global Ecology and Conservation, vol-23, year, PP. 1-13
- [6] Janos Katai, Agnes Olah Zsuposne, Magdolna Tallai, Tarek Alshaal, "
 Would fertilization history render the soil microbial communities and their activities
 more resistant to rainfall fluctuations?", Ecotoxicology and Environment Safety, ISSN:
 0147-6513 (Elsevier)
- [7] JánosKátai,ÁgnesOláhZsuposné,Magdolna Tállai,Tarek Alshaal, "Would fertilization history render the soil microbial communities and their activities more resistant to rainfall fluctuations?," Ecotoxicology and Environmental Safety, vol-201, 2020, PP. 1-11

- [8] Usman Ahmed, Jerry Chun-Wei Lin, Gautam Srivastava, Youcef Djenouri, "A nutrient recommendation system for soil fertilization based on Evolutionary Computation," Computers and Electronics in Agriculture, vol-189, 2021, PP. 1-7
- [9] A.Hussein⁷ Diogenes L. Antille, ShreevatsaKodur⁷ GuangnanChen,Jeff N.Tullberg, "Controlled traffic farming effects on productivity of grain sorghum, rainfall and fertilizer nitrogen use efficiency," Journal of Agriculture and Food Research, vol-3, 2021, PP. 1-17
- [10] Zujiao Shi, Donghua Liu, Miao Liu, Muhammad Bilal Hafeez, Pengfei Wen, Xiaoli Wang, Rui Wang, Xudong Zhang, Jun Li, "Optimized fertilizer recommendation method for nitrate residue control in a wheat–maize double cropping system in dryland farming," Field Crops Research, vol-271, 2021, PP. 1-10
- [11] Janmejay Pant, R.P. Pant, Manoj Kumar Singh, Devesh Pratap Singh, Himanshu Pant, "Analysis of agricultural crop yield prediction using statistical techniques of machine learning," Materials Today: Proceedings, vol-46, 2021, PP. 1-10
- [12] Benny Antony, "Prediction of the production of crops with respect to rainfall.," Environmental Research, vol-202, 2021, PP. 1-5
- [13] Akash Manish Lad, K. Mani Bharathi, B. Akash Saravanan, R. Karthik, "Factors affecting agriculture and estimation of crop yield using supervised learning algorithms," Materials Today: Proceedings, 2022, PP. 1-10
- [14] Raves Akhtar, Shabbir Ahmad Sofi, "Precision agriculture using IoT data analytics and machine learning," Journal of King Saud University Computer and Information Sciences, 2021, PP. 1-17
- [15] SaheedGarnaik, Prasanna Kumar Samant, Mitali Mandal, Tushar Ranjan Mohanty,

Sanat Kumar Dwibedi, Ranjan Kumar Patra, Kiran Kumar Mohapatra, R.H. Wanjari, DebadattaSethi, Dipaka Ranjan Sena, Tek Bahadur Sapkota, Jagmohan Nayak, Sridhar Patra, Chiter Mal Parihar, Hari Sankar Nayak, "Untangling the effect of soil quality on rice productivity under a 16-years long-term fertilizer experiment using conditional random forest," Computers and Electronics in Agriculture, vol-197,2022, PP. 1-10

[16] RubbyAworka, LontsiSaadio Cedric, Wilfried Yves Hamilton Adoni, JérémieThouakessehZoueu, Franck KalalaMutombo, Charles LebonMberiKimpolo, Tarik Nahhal, MoezKrichen, "Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries," Smart Agricultural Technology, vol-3, 2022, PP. 1-9

[17] Senthil Kumar Swami Durai, Mary DivyaShamili, "Smart farming using Machine Learning and Deep Learning techniques," Decision Analytics Journal, vol-2, 2022, PP. 1-30

[18] Aishwarya Daivagana, G gandhya, Bhumika Y, B Madhupriya, "An Efficient And Intelligent Decision Making For EcoFertilization" IJCRT Journal, ISSN;2320-2882

[19] Khan, Atiya; Vibhute, Amol D.; Mali, Shankar; Patil, C. H.," A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications" ECOLOGICAL INFORMATICS, JUL 2022

[20] Khaire, Utkarsh Mahadeo; Dhanalakshmi, R.," Stability Investigation of Improved Whale Optimization Algorithm in the Process of Feature Selection", IETE TECHNICAL REVIEW, MAR 4 2022

[21] Raut, Ekta R.; Bedmohata (Thakur), Monita A.; Chaudhari, Archana R." Comparative study of preparation and characterization of activated carbon obtained from sugarcane bagasse and rice husk by using H3PO4 and ZnCl2", MATERIALS TODAY-PROCEEDINGS,2022

[22] Golait, Manohar; Yadav, Manish Varun; Patil, Balasaheb H.; Baudha, Sudeep; Bramhane, Lokesh Kumar," Golait, Manohar; Yadav, Manish Varun; Patil, Balasaheb H.; Baudha, Sudeep; Bramhane, Lokesh Kumar", INTERNATIONAL JOURNAL OF MICROWAVE AND WIRELESS TECHNOLOGIES, OCT 2022

[23] Bramhane, Lokesh; Salankar, Suresh; Gaikwad, Mahendra; Panchore, Meena," Impact of Work Function Engineering in Charge Plasma Based Bipolar Devices", SILICON,JUN 2022

[24] Kumar, Rakesh; Panchore, Meena; Bramhane, Lokesh; Singh, Jawar" Comparative Performance and Reliability Analysis of Doping and Junction Free Devices with High-kappa/Vacuum Gate Dielectric", SILICON, JUN 2022

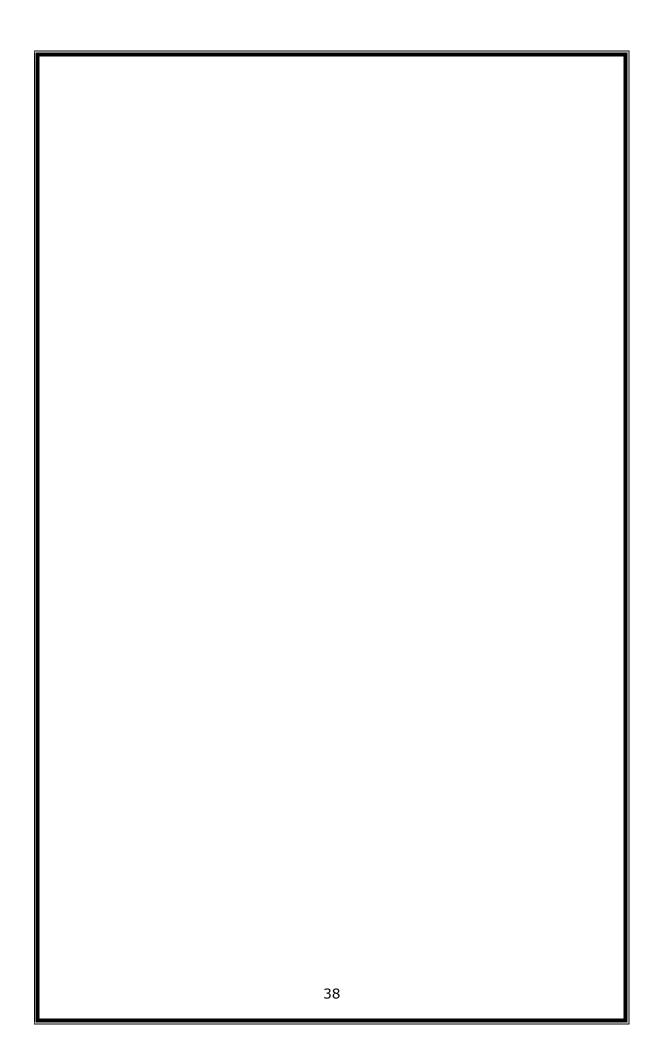
[25]Adjei, Seth A.; Baker, Ryan S.; Bahel, Vedant," Seven-Year Longitudinal Implications of Wheel Spinning and Productive Persistence", ARTIFICIAL INTELLIGENCE IN EDUCATION (AIED 2021), PT I,2021

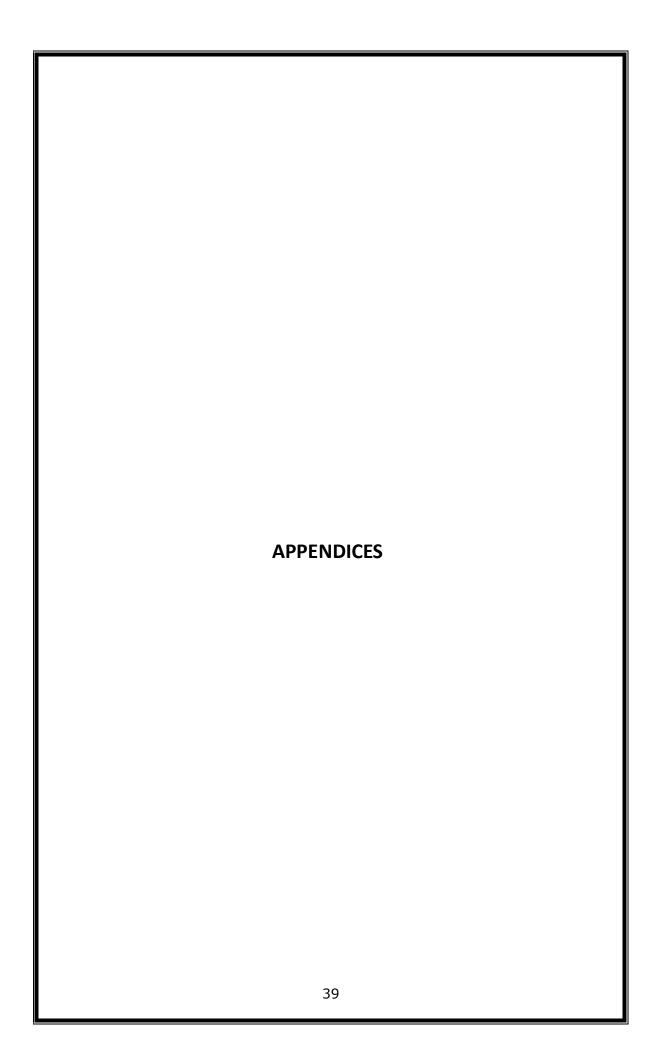
[26] Iwendi, Celestine; Bashir, Ali Kashif; Peshkar, Atharva; Sujatha, R.; Chatterjee, Jyotir Moy; Pasupuleti, Swetha; Mishra, Rishita; Pillai, Sofia; Jo, Ohyun," COVID-19 Patient Health Prediction Using Boosted Random Forest Algorithm", FRONTIERS IN PUBLIC HEALTH, JUL 3 2020

[27] Khekare, Ganesh; Verma, Shrutika; Sur, Samay; Haldkar, Rakesh; Moon, Pratik" Design of Fruit Segregation and Packaging Machine", 2020 INTERNATIONAL CONFERENCE ON COMPUTATIONAL PERFORMANCE EVALUATION (COMPE-2020),2020

[28] Dhule, Chetan; Shrawankar, Urmila," POF-SVLM: pareto optimized framework for seamless VM live migration", COMPUTING,OCT2020

- [29] Kumar, K.; Farande, K. U.; Ajai, S.; Sahu, T. K.; Raut, A.; Farande, S.; Tichkule, Y," Kumar, K.; Farande, K. U.; Ajai, S.; Sahu, T. K.; Raut, A.; Farande, S.; Tichkule, Y", TRENDS IN MANUFACTURING PROCESSES, ICFTMM 2018,2020
- [30] Kumari, Parveen; Singh, Sugandha; Aggarwal, Gaurav," Kumari, Parveen; Singh, Sugandha; Aggarwal, Gaurav", INTELLIGENT COMMUNICATION TECHNOLOGIES AND VIRTUAL MOBILE NETWORKS, ICICV 2019,2020
- [31] Yerne, Shubhangi; Shrawankar, Urmila," Analysis of Geographical Change Detection using Satellite Images", 2020 IEEE 9TH INTERNATIONAL CONFERENCE ON COMMUNICATION SYSTEMS AND NETWORK TECHNOLOGIES (CSNT 2020),2020
- [32] Dutta, Prateek; Sarkar, Abhiroop; Ambekar, Yash; Pek, Hui Ting; Juwono, F. H.; Sakarkar, Gopal," A Computational Approach for Predicting the Termination of COVID-19", 2022 INTERNATIONAL CONFERENCE ON GREEN ENERGY, COMPUTING AND SUSTAINABLE TECHNOLOGY (GECOST)2022
- [33] Mate, Nikita; Akre, Durva; Patil, Gaurav; Sakarkar, Gopal; Basuki, Thomas Anung," Mate, Nikita; Akre, Durva; Patil, Gaurav; Sakarkar, Gopal; Basuki, Thomas Anung"," Emotion Classification of Songs Using Deep Learning", 2022 INTERNATIONAL CONFERENCE ON GREEN ENERGY, COMPUTING AND SUSTAINABLE TECHNOLOGY (GECOST),2022
- [34] Mishra, Prerna; Kumar, Santosh; Chaube, Mithilesh Kumar; Shrawankar, Urmila," ChartVi: Charts summarizer for visually impaired", JOURNAL OF COMPUTER LANGUAGES, APR 2022
- [35] Talmale, Girish; Shrawankar, Urmila," Cluster formation techniques for hierarchical real time tasks allocation on multiprocessor system", CONCURRENCY AND COMPUTATION-PRACTICE & EXPERIENCE, NOV 10 2021





APPENDICES PATIENT STATUS

Ujjwal Sadani, Utkarsh Gomase, Vedant Dagwar, Varun Rewatkar, Sampada Wazalwar

AN ECO-FERTILIZATION SYSTEM

Application No. 202421020445

Date of Filing: 19/3/2024 Status: Filed

