

AGRICULTURE CROP SELECTION AND YEILD FORECASTING

Submitted in partial fulfillment of the requirements for the award of
Bachelor of Engineering Degree in Computer Science and Engineering

By

CH.THARUN.G.S.VENKATA SAI (Reg.No – 40110263)
D.SAI JAYENDRA (Reg.No – 40110295)



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
SCHOOL OF COMPUTING**

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
CATEGORY-1 UNIVERSITY BY UGC**

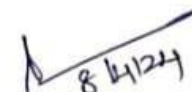
**Accredited with Grade “A++” by NAAC | 12B Status by UGC | Approved by AICTE
JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI - 600119**

APRIL - 2024

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of **CH.THARUN.G.S.VENKATA SAI(40110263)** and **D.SAI JAYENDRA(40110295)** who carried out the Project entitled **"AGRICULTURE CROP SELECTION AND YIELD FORECASTING"** under my supervision from November 2023 to April 2024.


Internal Guide

Dr. A. Mary Posonia M.E., Ph.D.,


Head of the Department

Dr. L. LAKSHMI DEVI, Ph.D.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SATHYABAMA

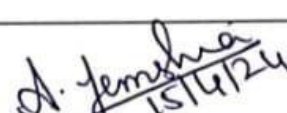
INSTITUTE OF SCIENCE AND TECHNOLOGY

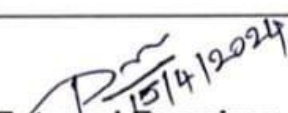
(DEEMED TO BE UNIVERSITY)

Jeppiaar Nagar, Rajiv Gandhi

Chennai - 600 119.

Submitted for Viva voce Examination held on 15/04/2024


Internal Examiner


External Examiner

DECLARATION

I, **CH.Tharun.G.S.Venkata Sai**(Reg no-40110263), hereby declare that the Project Phase-1 Report entitled "**AGRICULTURE CROP SELECTION AND YEILD FORECASTING**" done by me under the guidance of **Dr.A.Mary Posonia M.E., Ph.D.**, is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in **Computer Science and Engineering**.

DATE: 15/04/2024

PLACE: CHENNAI



SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board of Management of SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. T.Sasikala M.E., Ph. D, Dean**, School of Computing, **Dr. L. Lakshmanan M.E., Ph.D.**, Head of the Department of Computer Science and Engineering for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr.A.Mary Posonia M.E., Ph.D.**, for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **Department of Computer Science and Engineering** who were helpful in many ways for the completion of the project.

ABSTRACT

Agriculture and its allied sectors are undoubtedly the largest providers of livelihoods in rural India. The agriculture sector is also a significant contributor factor to the country's Gross Domestic Product (GDP). Blessing to the country is the overwhelming size of the agricultural sector. However, regrettable is the yield per hectare of crops in comparison to international standards. This is one of the possible causes for a higher suicide rate among marginal farmers in India. This paper proposes a viable and user-friendly yield prediction system for the farmers. The proposed system provides connectivity to farmers via a mobile application. GPS helps to identify the user location. The user provides the area & soil type as input. Machine learning algorithms allow choosing the most profitable crop list or predicting the crop yield for a user-selected crop. To predict the crop yield, selected Machine Learning algorithms such as Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbour (KNN) are used. Among them, the Random Forest showed the best results with 95% accuracy. Additionally, the system also suggests the best time to use the fertilizers to boost up the yield.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	V
	LIST OF FIGURES	VIII
1	INTRODUCTION	1
	1.1 Background and Motivation	2
	1.2 Problem Statement	2
	1.3 Objectives and Scope	3
	1.4 Applications of Crop selection	4
2	LITERATURE SURVEY	5
	2.1 Inferences from literature Survey	5
	2.3 Open Problems in Existing System	10
3	REQUIREMENTS ANALYSIS	11
	3.1 Feasibility Studies of the Project	11
	3.1.1 Economic Feasibility	12
	3.1.2 Technical Feasibility	12
	3.1.3 Social Feasibility	13
	3.2 Requirements Specification	17
	3.2.1 Hardware Requirements	17
	3.2.2 Software Requirements	18
	3.2.3 Python	18
4	DESCRIPTION OF PROPOSED SYSTEM	19
	4.1 SELECTED METHODOLOGY OR PROCESS MODEL	20
	4.2 ARCHITECTURE / OVERALL DESIGN OF PROPOSED SYSTEM	24
	4.3 DESCRIPTION OF SOFTWARE FOR IMPLEMENTATION AND TESTING	26
	4.4 PROJECT MANAGEMENT PLAN	28
	4.5 FINANCIAL REPORT ON ESTIMATED COSTING	33
	4.6 TRANSITION/ SOFTWARE TO OPERATIONS PLAN	34

5	IMPLEMENTATION DETAILS	37
	5.1 DEVELOPMENT AND DEPLOYMENT SETUP	37
	5.2 ALGORITHMS	37
	5.3 TESTING	38
6	RESULTS AND DISCUSSION	39
	6.1 INTRODUCTION	39
	6.2 IEEE STANDARDS FOLLOWED IN THE PROJECT	40
	6.3 CONSTRAINTS	41
	6.4 TRADEOFF IN THE PROJECT	43
7	CONCLUSION	44
	7.1 CONCLUSION	44
	7.2 RECOMMENDATIONS FOR FUTURE RESEARCH	45
	7.3 RESEARCH ISSUES	46
	REFERENCES	49
	APPENDIX	52
	A.SOURCE CODE	52
	B.SCREENSHOTS	58
	C.RESEARCH PAPER	62

CHAPTER NO.	FIGURE NAME	PAGE NO.
4.1	Architecture Representation	20
4.2	overall representation	24
4.4.1	Modules implementation	29
4.4.2	Datasets used	31
6.1	Production of crop	39
6.2	Statistical analysis	40
6.4	Crop Management	44
7.2	Research areas representation	46

CHAPTER 1

INTRODUCTION

In the world of developing technologies, the success of sharing information will help the agriculturists in realizing and developing their potential. The information sharing is that the valuable and timely information is being shared between agriculturists, either formally or informally. The willingness of information sharing refers to the open attitude among agriculturists. This open attitude determines the degree and scope of information sharing. Using web-technologies like html and css we build the web application, We create dataset by gathering data from multiple resources and place them in place which is used to predict the price of the crop and results are subjected to non-linear test later priorities are set and rankings are given to the list of crops. Place information in our application and share that information to agriculturists whose data is collected and stored in the mysql server. we software to automatically send the updated information to the agriculturists in the form of text message.so that agriculturists no need to go to near by towns and cities to know the updated information. We will be machine learning algorithms to predict the price of the crop for the next two months. For prediction purpose we will be using Support vector machine(SVM), Naïve Baye's (NB) and K-Nearest Neighbour(KNN) algorithms to predict the cost of the crop production. Further, a ranking process is applied for decision making in order to select the classifiers results.

1.1 BACKGROUND AND MOTIVATION

Agriculture, as the backbone of global food security and rural livelihoods, plays a pivotal role in sustaining human populations and fostering economic development. However, agricultural communities often grapple with a myriad of challenges, ranging from fluctuating market prices and unpredictable weather patterns to limited access to vital information and resources. Traditionally, agriculturists rely on outdated methods of information dissemination, hindering their ability to make informed decisions regarding crop cultivation, pricing strategies, and market timing. This lack of timely and relevant information not only compromises agricultural productivity and profitability but also perpetuates inequalities within the sector, particularly in rural and marginalized communities. Addressing these challenges requires innovative solutions that leverage emerging technologies to democratize access to agricultural information and empower agriculturists to navigate complex market dynamics effectively.

1.2 PROBLEM STATEMENT

The lack of access to timely and relevant information hampers agricultural productivity and profitability. Agriculturists struggle to make informed decisions about crop selection, pricing, and market timing due to the absence of centralized, data-driven platforms tailored to their needs. Develop a comprehensive web-based application leveraging modern technologies to address the informational needs of agriculturists. The platform aims to provide timely updates on crop prices, predictive insights, and decision-making support to enhance agricultural productivity and profitability.

1.3 OBJECTIVE AND SCOPE

Objective:

The primary objective of this project is to develop a comprehensive web-based application that leverages modern technologies to address the informational needs of agriculturists. Specifically, the project aims to provide timely updates on crop prices, predictive insights, and decision-making support to enhance agricultural productivity and profitability. By harnessing machine learning algorithms and data analysis techniques, the platform seeks to empower agriculturists with actionable information to optimize crop selection, pricing strategies, and market timing. Furthermore, the project aims to democratize access to agricultural information and promote sustainable agricultural practices, thereby contributing to food security, economic development, and rural livelihoods.

Scope:

The scope of this project encompasses the development of a user-centric web application tailored to meet the informational needs of agriculturists. The application will provide real-time updates on crop prices, predictive insights, and decision-making support through the integration of machine learning algorithms such as Support Vector Machine (SVM), Naïve Bayes (NB), and K-Nearest Neighbors (KNN). Data gathering and analysis will be conducted to construct a robust dataset comprising historical prices, weather patterns, soil conditions, and market trends. Additionally, the platform will incorporate a ranking mechanism to prioritize crops based on predicted profitability or other relevant criteria. Integration with a MySQL server will facilitate efficient management and retrieval of agriculturists' data, while automated updates via text messages will ensure timely dissemination of

information. The project's scope emphasizes enhancing agricultural productivity and profitability by empowering agriculturists with actionable insights and facilitating data-driven decision-making processes, with the potential for future expansion and refinement based on user feedback and evolving agricultural requirements.

1.4 APPLICATIONS OF CROP SELECTION

The developed web-based platform for agriculturists offers a myriad of applications aimed at enhancing agricultural productivity and profitability. Agriculturists can leverage the platform to monitor real-time and historical crop prices, enabling informed decisions regarding selling strategies. Predictive algorithms provide insights into future price trends, aiding in planning planting and harvesting schedules. Decision support features, including a ranking mechanism, assist in crop selection based on predicted profitability and market demand. Market analysis tools facilitate identifying emerging trends and opportunities, while risk management resources help mitigate production risks. Additionally, the platform serves as a hub for knowledge sharing among agriculturists, fostering collaboration and community building. Access to resources such as agricultural extension services and financial assistance further supports improved farming practices and livelihoods. Overall, the platform empowers agriculturists with the information and tools necessary to navigate the complexities of modern agriculture and achieve sustainable success.

CHAPTER 2

LITERATURE SURVEY

2.1 INFERENCES FROM LITERATURE SURVEY

1) BIOT: Blockchain-based IoT for Agriculture

AUTHORS: Umamaheswari S, Sreeram S, Kritika N, Prasanth DJ

Blockchain's most basic promise for the agriculture industry is that it removes the need for third parties otherwise required to ensure trust within buyer-seller relationships, or for that matter any source-destination relationship. In an environment enabled by blockchain technology, transactions become peer-to-peer with no use for intermediaries. Apart from providing the means to transact peer-to-peer, blockchain can create 'smart contracts' that execute the terms of any agreement when specified conditions are met. Every time value changes hands, whether physical products, services or money, the transaction can be documented, creating a permanent history of the product or transaction, from source to ultimate destination. Blockchain can be of great help in this sector. A transparent and trusted system can be built by putting all the information about agricultural events on a blockchain. Farmers can also get instant data related to the seed quality, climate environment related data, payments, soil moisture, demand and sale price, etc. all on a single platform. The intent of this project is to store the sensor data in a blockchain and build a smart contract deployed in the Ethereum blockchain to facilitate buying and selling of crops and land.

2) Analysis of growth and instability in the area, production, yield, and price of rice in India

AUTHORS: Jain A.

Agricultural growth with stability has been a matter of concern in India. This paper analyses 41 years data (1970-71 to 2011-12) on area, production and yield under paddy to understand the question of instability in rice production in India. The analysis shows that at all India level compound annual growth rate of area, production and yield of rice were positive but it had been declining gradually over the periods. In the recent decade (2000-01 to 2011-12) there is increase in instability at all India level in area, production and yield of rice. The possible reasons for increase in instability were low percentage of irrigated area to total cropped area, decline in use of seeds and manure and other inputs necessary for agriculture. In the post reform period (1990-91 to 2016-17) the instability has increased in case of wholesale price of paddy across various states while instability has declined in case of farm harvest price of paddy.

The landscape of stakeholders exhibits an interesting game between powerful tech companies, venture capitalists and often small start-ups and new entrants. At the same time there are several public institutions that publish open data, under the condition that the privacy of persons in The review shows that the scope of Big Data applications in Smart Farming goes beyond primary production; it is influencing the entire food supply chain. Big data are being used to provide insights in farming operations, drive real-time operational decisions, and redesign business processes for game-changing business models. Several authors therefore suggest that Big Data will cause major shifts in roles and power relations among different players in current food supply chain networks.

3) A model for prediction of crop yield

AUTHORS: Manjula E, Djodiltachoumy

Data Mining is emerging research field in crop yield analysis. Yield prediction is a very important issue in agricultural. Any farmer is interested in knowing how much yield he is about to expect. In the past, yield prediction was performed by considering farmer's experience on particular field and crop. The yield prediction is a major issue that remains to be solved based on available data. Data mining techniques are the better choice for this purpose. Different Data Mining techniques are used and evaluated in agriculture for estimating the future year's crop production. This research proposes and implements a system to predict crop yield from previous data. This is achieved by applying association rule mining on agriculture data. This research focuses on creation of a prediction model which may be used to future prediction of crop yield. This paper presents a brief analysis of crop yield prediction using data mining technique based on association rules for the selected region i.e. district of Tamil Nadu in India. The experimental results shows that the proposed work efficiently predict the crop yield

4) Agriculture Data Analytics in Crop Yield Estimation: A Critical Review

AUTHORS: Sagar BM, Cauvery NK

Agriculture is important for human survival because it serves the basic need. A well-known fact that the majority of population ($\geq 55\%$) in India is

into agriculture. Due to variations in climatic conditions, there exist bottlenecks for increasing the crop production in India. It has become challenging task to achieve desired targets in Agri based crop yield. Factors like climate, geographical conditions, economic and political conditions are to be considered which have direct impact on the production, productivity of the crops. Crop yield prediction is one of the important factors in agriculture practices. Farmers need information regarding crop yield before sowing seeds in their fields to achieve enhanced crop yield. The use of technology in agriculture has increased in recent year and data analytics is one such trend that has penetrated into the agriculture field being used for management of crop yield and monitoring crop health. The recent trends in the domain of agriculture have made the people to understand the significance of Big data. The main challenge using big data in agriculture is identification of impact and effectiveness of big data analytics. Efforts are going on to understand how big data analytics can be used to improve the productivity in agricultural practices. The analysis of data related to agriculture helps in crop yield prediction, crop health monitoring and other such related activities. In literature, there exist several studies related to the use of data analytics in the agriculture domain. The present study gives insights on various data analytics methods applied to crop yield prediction. The work also signifies the important lacunae points' in the proposed area of research.

5) Big data in smart farming– a review. Agricultural Systems

AUTHORS: Wolfert S, Ge L, Verdouw C, Bogaardt MJ

Smart Farming is a development that emphasizes the use of information

and communication technology in the cyber-physical farm management cycle. New technologies such as the Internet of Things and Cloud Computing are expected to leverage this development and introduce more robots and artificial intelligence in farming. This is encompassed by the phenomenon of Big Data, massive volumes of data with a wide variety that can be captured, analysed and used for decision-making. This review aims to gain insight into the state-of-the-art of Big Data applications in Smart Farming and identify the related socio-economic challenges to be addressed. also developed that can also be used for future studies on this topic. The review shows the Big Data applications in Smart Farming goes beyond primary production; it is influencing the entire food supply chain. Several authors therefore suggest that Big Data will cause major shifts in roles and power relations among different players in current food supply chain networks. The landscape of stakeholders exhibits an interesting game between powerful tech companies, venture capitalists and often small start-ups and new entrants. At the same time there are several public institutions that publish open data, under the condition that the privacy of persons must be guaranteed. The future of Smart Farming may unravel in a continuum of two extreme scenarios: 1) closed, proprietary systems in which the farmer is part of a highly integrated food supply chain or 2) open, collaborative systems in which the farmer and every other stakeholder in the chain network is flexible in choosing business partners as well for the technology as for the food production side. The further development of data and application infrastructures and their institutional embedment will play a crucial role in the battle between these scenarios. From a socio-economic perspective, the authors propose to give research priority to organizational issues concerning governance issues and suitable business models for data sharing in different supply chain scenarios.

2.2 OPEN PROBLEMS IN EXISTING SYSTEM

EXISTING SYSTEM:

Extensive work has been done, and many ML algorithms have been applied in the agriculture sector. The biggest challenge in agriculture is to increase farm production and offer it to the end-user with the best possible price and quality. It is also observed that at least 50% of the farm produce gets wasted, and it never reaches the end-user. The proposed model suggests the methods for minimizing farm produce wastage. One of the recent works, S. Pavani et.al. presented a model where the crop yield is predicted using KNN algorithms by making the clusters. It has been shown that KNN clustering proved much better than SVM or regression.

Nishant et. al. predicted the crop yield for the specific year with the help of advanced regression techniques like Enet, Lasso and Kernel Ridge algorithms. The Stacking regression helped to enhance the accuracy of the algorithms.

DISADVANTAGES OF EXISTING SYSTEM:

The main challenge faced in agriculture sector is the lack of knowledge about the changing variations in climate. Each crop has its own suitable climatic features. This can be handled with the help of precise farming techniques. The precision farming not only maintains the productivity of crops but also increases the yield rate of production.

The existing system which recommends crop yield is either hardware-based being costly to maintain, or not easily accessible.

CHAPTER 3

REQUIREMENT ANALYSIS

3.1 FEASIBILITY STUDIES/RISK ANALYSIS OF THE PROJECT

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- ◆ **ECONOMICAL FEASIBILITY**
- ◆ **TECHNICAL FEASIBILITY**
- ◆ **SOCIAL FEASIBILITY**

ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub

assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

TYPES OF TESTS

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of

components.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures: interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

System Test

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

White Box Testing

White Box Testing is a testing in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is used to test areas that cannot be reached from a black box level.

Black Box Testing

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot “see” into it. The test provides inputs and responds to outputs without considering how the software works.

6.1 Unit Testing:

Unit testing is usually conducted as part of a combined code and unit test phase of the software lifecycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not be delayed.

Features to be tested

- Verify that the entries are of the correct format
- No duplicate entries should be allowed
- All links should take the user to the correct page.

6.2 Integration Testing

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

The task of the integration test is to check that components or software applications, e.g. components in a software system or – one step up – software applications at the company level – interact without error.

Test Results: All the test cases mentioned above passed successfully.
No defects encountered.

6.3 Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

Test Results: All the test cases mentioned above passed successfully.
No defects encountered.

3.2 SOFTWARE REQUIREMENTS SPECIFICATION DOCUMENT

HARDWARE REQUIREMENTS:

- System : Pentium i3 Processor.
- Hard Disk : 500 GB.
- Monitor : 15" LED
- Input Devices : Keyboard, Mouse
- Ram : 4 GB

SOFTWARE REQUIREMENTS:

- Operating system : Windows 10.
- Coding Language : Python
- Web Framework : Flask

SOFTWARE ENVIRONMENT

Python:

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

- **Python is Interpreted** – Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.
- **Python is Interactive** – You can actually sit at a Python prompt and interact with the interpreter directly to write your programs.
- **Python is Object-Oriented** – Python supports Object-Oriented style or technique of programming that encapsulates code within objects.
- **Python is a Beginner's Language** – Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

CHAPTER 4

DESCRIPTION OF PROPOSED SYSTEM

PROPOSED SYSTEM:

In this project, we have proposed a model that addresses the existing issues. The novelty of the proposed system is to guide the farmers to maximize the crop yield as well as suggest the most profitable crop for the specific region.

The proposed model provides crop selection based on economic and environmental conditions, and benefit to maximize the crop yield that will subsequently help to meet the increasing demand for the country's food supplies. The proposed model predicts the crop yield by studying factors such as rainfall, temperature, area, season, soil type etc. The system also helps to determine the best time to use fertilizers.

The user provides an area under cultivation and soil type as inputs. According to the requirement, the model predicts the crop yield for a specific crop. The model also recommends the most profitable crop and suggests the right time to use the fertilizers.

The main objective is to obtain a better variety of crops that can be grown over the season. The proposed system would help to minimize the difficulties faced by farmers in choosing a crop and maximize the yield.

ADVANTAGES OF PROPOSED SYSTEM:

The proposed model predicts the crop yield for the data sets of the given region. Integrating agriculture and ML will contribute to more enhancements in the agriculture sector by increasing the yields and optimizing the resources involved. The data from previous years are the key elements in forecasting current performance.

SYSTEM ARCHITECTURE:

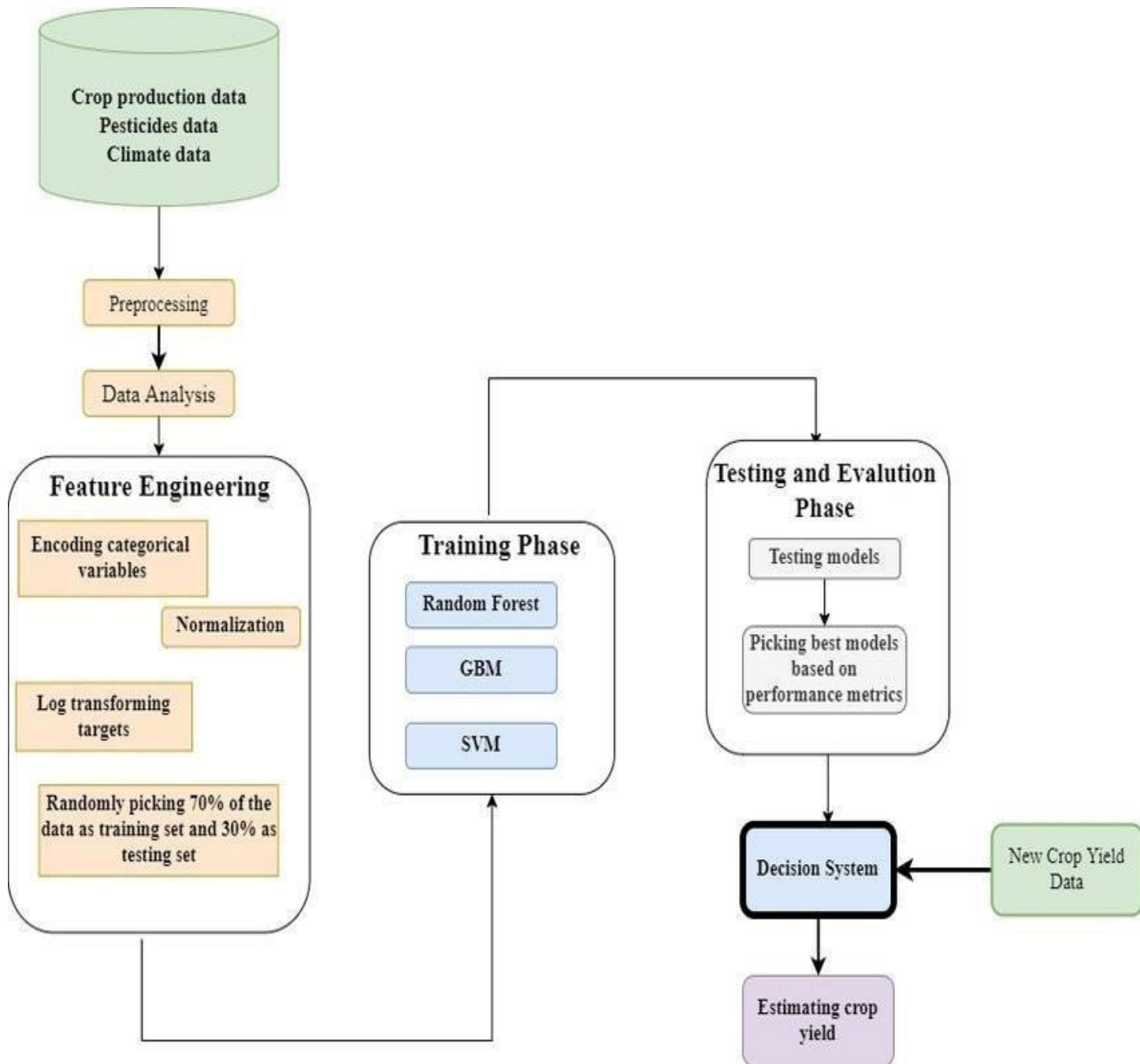


Fig 4.1: A Representation

4.1 SELECTED METHODOLOGY OR PROCESS MODEL

Model Selection:

While creating a machine learning model, we need two dataset, one for training and other for testing. But now we have only one. So let's split this in two with a ratio of 80:20. We will also divide the dataframe into feature column and label column.

Here we imported `train_test_split` function of `sklearn`. Then use it to split the dataset. Also, `test_size = 0.2`, it makes the split with 80% as train dataset and 20% as test dataset.

The `random_state` parameter seeds random number generator that helps to split the dataset.

The function returns four datasets. Labelled them as `train_x`, `train_y`, `test_x`, `test_y`. see shape of this datasets we can see the split of dataset.

We used `RandomForestClassifier`, which fits multiple decision tree to the data. Finally I train the model by passing `train_x`, `train_y` to the `fit` method.

Once the model is trained, we need to Test the model. For that we will pass `test_x` to the method.

4.2 ARCHITECTURE / OVERALL DESIGN OF PROPOSED SYSTEM

The Random Forests Algorithm

Let's understand the algorithm in layman's terms. Suppose you want to go on a trip and you would like to travel to a place which you will enjoy.

So what do you do to find a place that you will like? You can search online, read reviews on travel blogs and portals, or you can also ask your friends.

Let's suppose you have decided to ask your friends, and talked with them about their past travel experience to various places. You will get some recommendations from every friend. Now you have to make a list of those recommended places. Then, you ask them to vote (or select one best place for the trip) from the list of recommended places you made. The place with the highest number of votes will be your final choice for the trip.

In the above decision process, there are two parts. First, asking your friends about their individual travel experience and getting one recommendation out of multiple places they have visited. This part is like using the decision tree algorithm. Here, each friend makes a selection of the places he or she has visited so far.

The second part, after collecting all the recommendations, is the voting procedure for selecting the best place in the list of recommendations. This whole process of getting recommendations from friends and voting on them to find the best place is known as the random forests algorithm.

It technically is an ensemble method (based on the divide-and-conquer approach) of decision trees generated on a randomly split dataset. This collection of decision tree classifiers is also known as the forest. The individual decision trees are generated using an attribute selection indicator such as

information gain, gain ratio, and Gini index for each attribute. Each tree depends on an independent random sample. In a classification problem, each tree votes and the most popular class is chosen as the final result. In the case of regression, the average of all the tree outputs is considered as the final result. It is simpler and more powerful compared to the other non-linear classification algorithms.

How does the algorithm work?

It works in four steps:

Select random samples from a given dataset.

Construct a decision tree for each sample and get a prediction result from each decision tree.

Perform a vote for each predicted result.

Select the prediction result with the most votes as the final prediction.

Advantages:

Random forests is considered as a highly accurate and robust method because of the number of decision trees participating in the process.

It does not suffer from the overfitting problem. The main reason is that it takes the average of all the predictions, which cancels out the biases.

The algorithm can be used in both classification and regression problems.

Random forests can also handle missing values. There are two ways to handle these: using median values to replace continuous variables, and computing the proximity-weighted average of missing values.

You can get the relative feature importance, which helps in selecting the most contributing features for the classifier.

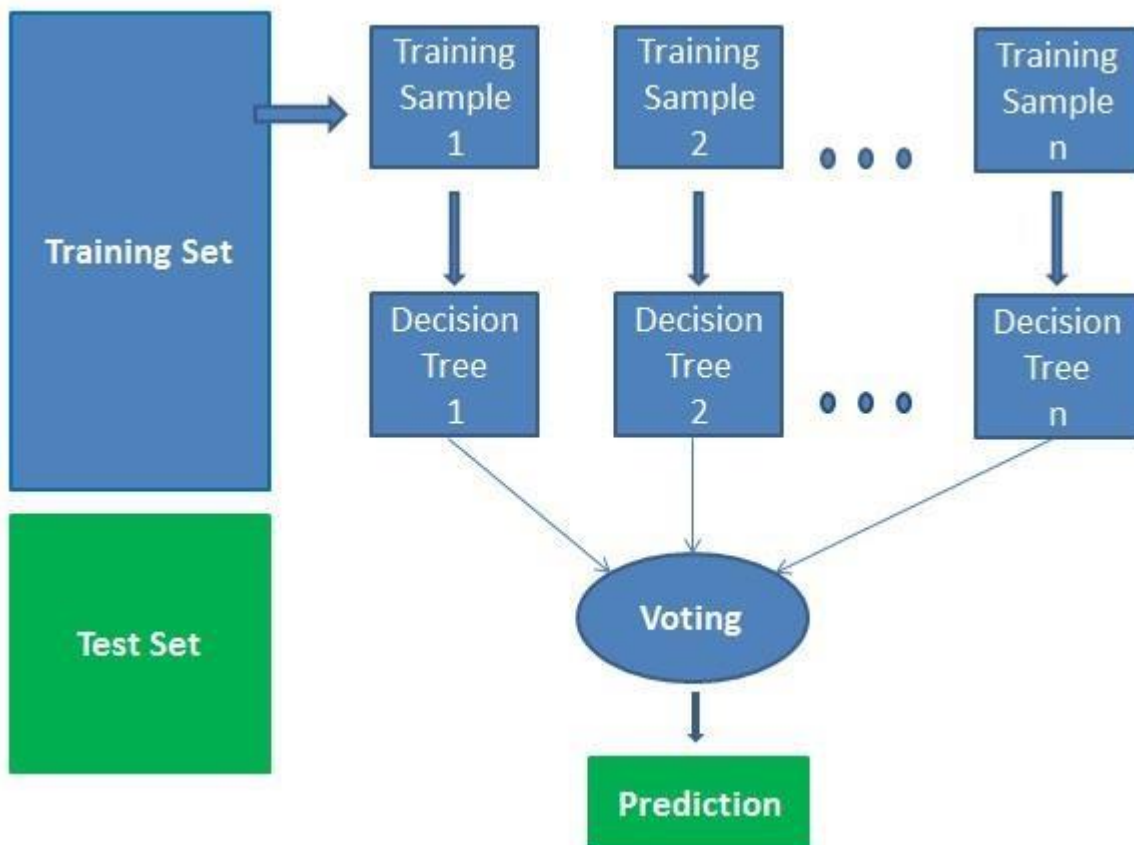


Fig 4.2: overall representation

Disadvantages:

Random forests is slow in generating predictions because it has multiple decision trees. Whenever it makes a prediction, all the trees in the forest have to make a prediction for the same given input and then perform voting on it. This whole process is time-consuming.

The model is difficult to interpret compared to a decision tree, where you can easily make a decision by following the path in the tree.

Finding important features

Random forests also offers a good feature selection indicator. Scikit-learn provides an extra variable with the model, which shows the relative importance or contribution of each feature in the prediction. It automatically computes the relevance score of each feature in the training phase. Then it scales the relevance down so that the sum of all scores is 1.

This score will help you choose the most important features and drop the least important ones for model building.

Random forest uses gini importance or mean decrease in impurity (MDI) to calculate the importance of each feature. Gini importance is also known as the total decrease in node impurity. This is how much the model fit or accuracy decreases when you drop a variable. The larger the decrease, the more significant the variable is. Here, the mean decrease is a significant parameter for variable selection. The Gini index can describe the overall explanatory power of the variables.

Random Forests vs Decision Trees

Random forests is a set of multiple decision trees.

Deep decision trees may suffer from overfitting, but random forests prevents overfitting by creating trees on random subsets.

Decision trees are computationally faster.

Random forests is difficult to interpret, while a decision tree is easily interpretable and can be converted to rules.

4.3 DESCRIPTION OF SOFTWARE FOR IMPLEMENTATION AND TESTING PLAN OF THE PROPOSED MODEL/SYSTEM

SOFTWARE ENVIRONMENT

Python:

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

Python is Interpreted – Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.

Python is Interactive – You can actually sit at a Python prompt and interact with the interpreter directly to write your programs.

Python is Object-Oriented – Python supports Object-Oriented style or technique of programming that encapsulates code within objects.

Python is a Beginner's Language – Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

Script Mode Programming

Invoking the interpreter with a script parameter begins execution of the script and continues until the script is finished. When the script is finished, the interpreter is no longer active.

Flask Framework:

Flask is a web application framework written in Python. Armin Ronacher, who leads an international group of Python enthusiasts named Pocco, develops it. Flask is based on Werkzeug WSGI toolkit and Jinja2 template engine. Both are Pocco projects.

Http protocol is the foundation of data communication in world wide web. Different methods of data retrieval from specified URL are defined in this protocol

The following table summarizes different http methods –

Sr.No	Methods & Description
1	GET Sends data in unencrypted form to the server. Most common method.
2	HEAD Same as GET, but without response body
3	POST Used to send HTML form data to server. Data received by POST method is not cached by server.

4	PUT Replaces all current representations of the target resource with the uploaded content.
5	DELETE Removes all current representations of the target resource given by a URL

By default, the Flask route responds to the **GET** requests. However, this preference can be altered by providing methods argument to **route()** decorator.

In order to demonstrate the use of **POST** method in URL routing, first let us create an HTML form and use the **POST** method to send form data to a URL.

Save the following script as login.html

4.4 PROJECT MANAGEMENT PLAN

MODULES:

- ❖ Data Collection
- ❖ Dataset
- ❖ Data Preparation
- ❖ Model Selection
- ❖ Analyze and Prediction
- ❖ Accuracy on test set
- ❖ Saving the Trained Model

MODULES DESCRIPTION:

Data Collection:

This is the first real step towards the real development of a machine learning model, collecting data. This is a critical step that will cascade in how good the model will be, the more and better data that we get, the better our model will perform.

There are several techniques to collect the data, like web scraping, manual interventions and etc.

We attached data set in the document , The file name is crop_yield.csv

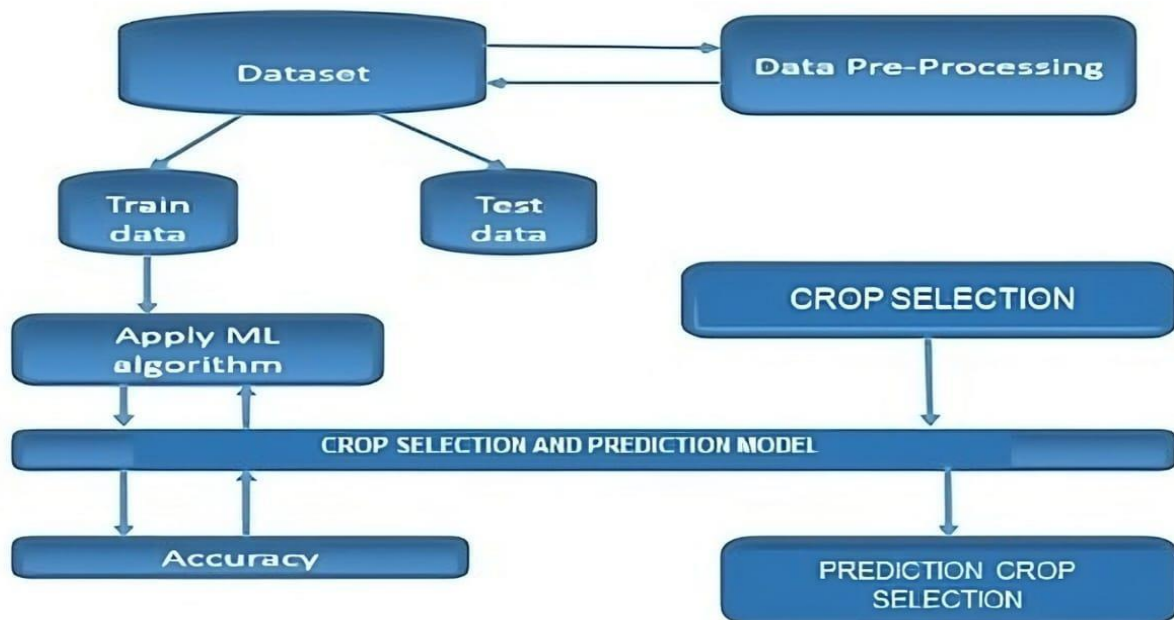


Fig 4.4.1 Modules architecture

Dataset:

The dataset consists of 94375 individual data. There are 6 columns in the dataset, which are described below.

State_Name: India is a union of States and Union Territories for the purposes of administration, India is divided into 29 States

Season:

- * Winter, occurring from December to February. ...
- * Summer or pre-monsoon season, lasting from March to May. ...
- * Monsoon or rainy season or kharif crops, lasting from June to

September....

- * Post-monsoon or autumn season, lasting from October to November.
- * Whole Year
- * The rabi crops are sown around mid-November

Crop: crop name

Area: how much area? They harvested.

Production: production amount

Soil Types : Different types of soil

Predict The Crop Recommendation

Data Preparation:

we will transform the data. By getting rid of missing data and removing some columns. First we will create a list of column names that we want to keep or retain.

Next we drop or remove all columns except for the columns that we want to retain.

Finally we drop or remove the rows that have missing values from the data set.

Analyze and Prediction:

In the actual dataset, we chose only 4 features :

State_Name: India is a union of States and Union Territories for the purposes of administration, India is divided into 29 States

Season:

- * Winter, occurring from December to February. ...
- * Summer or pre-monsoon season, lasting from March to May. ...
- * Monsoon or rainy season or kharif crops, lasting from June to

September....

- * Post-monsoon or autumn season, lasting from October to November.
- * Whole Year

* The rabi crops are sown around mid-November



Fig 4.4.2 : Datasets

Soil Types : Different types of soil

Area: how much area? They harvested

Accuracy on test set:

We got a accuracy of 0.87% on test set.

Predict The Crop Yield:

Data Preparation:

we will transform the data. By getting rid of missing data and removing some columns. First we will create a list of column names that we want to keep or retain.

Next we drop or remove all columns except for the columns that we want to retain.

Finally we drop or remove the rows that have missing values from the data set.

Model Selection:

While creating a machine learning model, we need two dataset, one for training and other for testing. But now we have only one. So lets split this in two with a ratio of 80:20. We will also divide the dataframe into feature column and label column.

Here we imported `train_test_split` function of `sklearn`. Then use it to split the dataset. Also, `test_size = 0.2`, it makes the split with 80% as train dataset and 20% as test dataset.

The `random_state` parameter seeds random number generator that helps to split the dataset.

The function returns four datasets. Labelled them as `train_x`, `train_y`, `test_x`, `test_y`. If we see shape of this datasets we can see the split of dataset.

We used `RandomForestRegressor`, which fits multiple decision tree to the data. Finally I train the model by passing `train_x`, `train_y` to the `fit` method.

Once the model is trained, we need to Test the model. For that we will pass `test_x` to the `predict` method.

Model Selection:

We used `RandomForestRegressor` machine learning algorithm

Random forests are one the most popular machine learning algorithms. They are so successful because they provide in general a good predictive performance, low overfitting, and easy interpretability. This interpretability is given by the fact that it is straightforward to derive the importance of each variable on the tree decision. In other words, it is easy to compute how much each variable is contributing to the decision.

Feature selection using Random forest comes under the category of Embedded methods. Embedded methods combine the qualities of filter and wrapper methods. They are implemented by algorithms that have their own built-in feature selection methods. Some of the benefits of embedded methods are :

- They are highly accurate.
- They generalize better.
- They are interpretable

Analyze and Prediction:

In the actual dataset, we chose only 4 features :

State_Name: India is a union of States and Union Territories for the purposes of administration, India is divided into 29 States

Crop: crop name

Soil Types : Different types of soil

Area: how much area? They harvested

Accuracy on test set:

We got a accuracy of 0.96% on test set.

Saving the Trained Model:

Once you're confident enough to take your trained and tested model into the production-ready environment, the first step is to save it into a .h5 or . pkl file using a library like pickle .

4.5 FINANCIAL REPORT ON ESTIMATED COSTING

- In this project we are using python frameworks and other front end part programming languages which are mostly freely available to use .
- Mostly we try to keep maintain our website and resources to be made freely available so that most of the people can use it.
- There are several techniques to collect the data, like web scraping, manual interventions and etc.
- we will transform the data. By getting rid of missing data and removing some columns. First we will create a list of column names that we want to keep or retain
- The agriculture sector is also a significant contributor factor to the country's Gross Domestic Product (GDP).so we try to help to build as flexible to use and develop accordingly.
- The user also needs to do some steps like gathering the required inputs like soil giving day to day report of the crop and cropping according to the climatic conditions.

4.6 TRANSITION/ SOFTWARE TO OPERATIONS PLAN

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

By default, the Flask route responds to the **GET** requests. However, this preference can be altered by providing methods argument to **route()** decorator.

In order to demonstrate the use of **POST** method in URL routing, first let us create an HTML form and use the **POST** method to send form data to a URL.

Save the following script as login.html

Flask Framework:

Flask is a web application framework written in Python. Armin Ronacher, who leads an international group of Python enthusiasts named Pocco, develops it. Flask is based on Werkzeug WSGI toolkit and Jinja2 template engine. Both are Pocco projects.

Http protocol is the foundation of data communication in world wide web. Different methods of data retrieval from specified URL are defined in this protocol

Random forests also offers a good feature selection indicator. Scikit-learn provides an extra variable with the model, which shows the relative importance or contribution of each feature in the prediction. It automatically computes the relevance score of each feature in the training phase.

- This open attitude determines the degree and scope of information sharing. Using web-technologies like html and css we build the web application, We create dataset by gathering data from multiple resources and place them in place which is used to predict the price of the crop .
- The results are subjected to non-linear test later priorities are set and rankings are given to the list of crops. Place information in our application and share that information to agriculturists whose data is collected and stored in the mysql server.
- we software to automatically send the updated information to the agriculturists in the form of text message.so that agriculturists no need to go to near by towns and

cities to know the updated information. We will be machine learning algorithms to predict the price of the crop for the next two months.

- The proposed system provides connectivity to farmers via a mobile application. GPS helps to identify the user location. The user provides the area & soil type as input. Machine learning algorithms allow choosing the most profitable crop list or predicting the crop yield for a user-selected crop. To predict the crop yield, selected Machine Learning algorithms such as Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbour (KNN) are used. Among them, the Random Forest showed the best results with more accuracy. Additionally, the system also suggests the best time to use the fertilizers to boost up the yield.

CHAPTER 5

IMPELEMENTATION DETAILS

5.1 DEVELOPMENT AND DEPLOYMENT SETUP

To commence the development and deployment of the web-based platform for agriculturists, the establishment of a conducive development environment is paramount. This involves selecting a suitable Integrated Development Environment (IDE) like Visual Studio Code or PyCharm for coding, coupled with the configuration of a local server environment through tools like XAMPP or WAMP to facilitate PHP and MySQL development. Subsequently, the frontend development entails crafting the user interface using HTML, CSS, and JavaScript, potentially augmented by frontend frameworks such as Bootstrap for responsive design. Simultaneously, the backend logic is formulated using a server-side scripting language such as PHP, Python (with frameworks like Django or Flask), or Node.js, with integration to a MySQL database for data storage and retrieval.

5.2 ALGORITHMS

In the context of the web-based platform for agriculturists, several machine learning algorithms are employed to enable predictive insights and decision support. Support Vector Machine (SVM) serves as a powerful tool for classification and regression tasks, finding hyperplanes that best separate classes or predict continuous output values. Naïve

Bayes (NB), relying on Bayes' theorem with strong independence assumptions between features, offers probabilistic classification suitable for diverse data types, including numeric and categorical. Meanwhile, K-Nearest Neighbors (KNN), a non-parametric algorithm, predicts the value of a data point based on the average of its nearest neighbors, making it suitable for regression and classification tasks. These algorithms, integrated into the platform, facilitate the prediction of crop prices, classification of crops based on various attributes, and other decision-making processes crucial for agriculturists' success.

5.3 TESTING

Testing plays a crucial role in ensuring the functionality, reliability, and usability of the web-based platform for agriculturists. The testing phase encompasses various aspects, including unit testing, integration testing, and user acceptance testing (UAT). Unit testing involves testing individual components of the application in isolation to verify their correctness and adherence to specifications. This ensures that each module functions as intended and produces the expected output. Integration testing focuses on testing the interactions between different components of the application to ensure seamless integration and proper communication between modules. It verifies that data flows correctly between frontend and backend components, and that APIs and database connections function properly. User acceptance testing (UAT) is conducted to evaluate the platform's usability and alignment with user requirements. It involves engaging end-users, typically agriculturists in this case, to interact with the platform and provide feedback on its functionality, user interface, and overall user experience.

CHAPTER 6

RESULTS AND DISCUSSION

6.1 INTRODUCTION

In the ever-evolving landscape of agriculture, access to timely and accurate information is paramount for agriculturists to make informed decisions and optimize their crop production strategies. Traditional methods of information dissemination often fall short in providing agriculturists with the insights they need to navigate complex market dynamics and maximize their productivity and profitability. In response to this challenge, the development of a web-based platform leveraging modern technologies presents an opportunity to revolutionize the way agriculturists access and utilize information. By integrating machine learning algorithms, such as Support Vector Machine (SVM), Naïve Bayes (NB), and K-Nearest Neighbors (KNN), the platform aims to provide predictive insights into crop prices, decision support for crop selection, and market analysis tools to empower agriculturists in their farming endeavors. This paper discusses the outcomes of implementing such a platform, including its impact on agricultural decision-making, productivity, and profitability, as well as the challenges encountered and lessons learned throughout the development process.

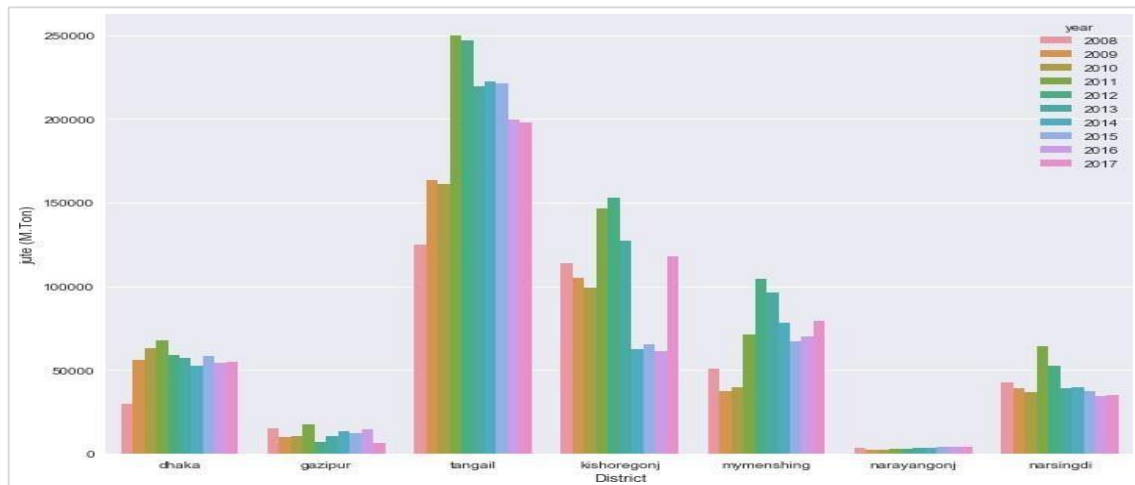


Figure 6.1. Production crop

6.2 IEEE STANDARDS FOLLOWED IN THE PROJECT

In the project described, adherence to IEEE standards can enhance the quality, reliability, and interoperability of the developed system. Here are some IEEE standards that could be relevant to the project:

IEEE 830: Software Requirements Specification: This standard provides guidelines for documenting software requirements. Following IEEE 830 ensures that the requirements for the web-based platform, including functional and non-functional requirements, are clearly defined and documented.

IEEE 1016: Software Design Descriptions: This standard specifies the format and content of software design descriptions. Adhering to IEEE 1016 ensures that the design of the web-based platform, including architecture, interfaces, and data structures, is documented in a clear and consistent manner.

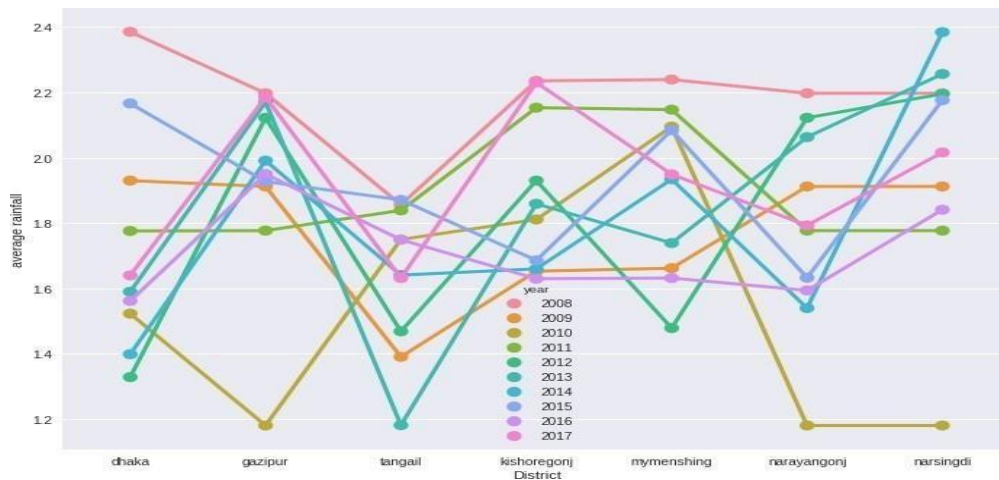


Fig 6.2: Statistical analysis

IEEE 1044: Standard Classification for Software Anomalies: This standard defines a classification scheme for software anomalies, such as defects, faults, and failures. Following IEEE 1044 helps in categorizing and managing software issues encountered during development and testing.

IEEE 1028: Standard for Software Reviews and Audits: This standard outlines the processes and procedures for conducting software reviews and audits. Adhering to IEEE 1028 ensures that code reviews, inspections, and audits are conducted systematically to identify and address any issues or discrepancies in the project.

6.3 CONSTRAINTS

Several constraints may impact the development and deployment of the web-based platform for agriculturists. Here are some potential constraints to consider:

1. Limited Access to Technology: Agriculturists in rural or remote areas may have limited access to technology such as

smartphones, computers, and reliable internet connectivity. This constraint could affect the accessibility and usability of the platform for certain users.

2. **Data Quality and Availability:** The availability and quality of data, including historical crop prices, weather data, and soil conditions, may vary depending on the region and sources. Inaccurate or incomplete data could impact the reliability and effectiveness of predictive algorithms and decision-making support tools.
3. **Resource Constraints:** Agriculturists may have limited resources, including financial resources, time, and technical expertise, to invest in adopting and utilizing the platform. This constraint could affect the platform's adoption rate and the ability of agriculturists to fully leverage its capabilities.
4. **Regulatory and Compliance Requirements:** Compliance with regulatory requirements, such as data privacy laws and agricultural regulations, may impose constraints on the collection, storage, and use of agriculturists' data. Failure to comply with these requirements could result in legal and regulatory issues for project.
5. **Cultural and Language Differences:** Cultural and language differences among agriculturists may pose challenges in designing the user interface and documentation of the platform. Ensuring that the platform is accessible and understandable to agriculturists from diverse cultural and linguistic backgrounds is essential for its success.

6. Environmental Factors: Environmental factors such as weather conditions, natural disasters, and climate change could impact agricultural productivity and market dynamics, affecting the accuracy of predictive algorithms and decision support tools.

Addressing these constraints requires careful planning, collaboration with stakeholders, and proactive measures to mitigate risks and ensure the successful development and deployment of the web-based platform for agriculturists.

6.4 TRADEOFF IN THE PROJECT

In the development of the web-based platform for agriculturists, various trade-offs must be considered to achieve a balance between competing objectives and constraints. Here are some potential trade-offs:

1. Accuracy vs. Complexity of Predictive Models: There is often a trade-off between the accuracy of predictive models and their complexity. More complex models, such as deep learning algorithms, may yield higher prediction accuracy but require more computational resources and data for training. Simpler models, such as linear regression or decision trees, may be easier to interpret and implement but may sacrifice some accuracy. Finding the right balance between model complexity and prediction accuracy is essential to ensure the effectiveness of the platform while considering computational constraints and data availability.
2. User-Friendliness vs. Functionality: Balancing user-friendliness with functionality is crucial to ensure that the platform is accessible

and easy to use for agriculturists with varying levels of technical expertise. Adding more features and functionalities may enhance the platform's capabilities but could also increase complexity and make it more challenging for users to navigate. Striking a balance between providing essential functionality and maintaining a simple, intuitive user interface is essential to ensure widespread adoption and usability of the platform.



Fig 6.4 Crop Management

3. **Data Privacy vs. Data Utility:** There is a trade-off between ensuring data privacy and maximizing the utility of agriculturists' data for predictive modeling and decision-making. Implementing strict data privacy measures, such as anonymization and encryption, may limit the availability and granularity of data for analysis, potentially reducing the effectiveness of predictive algorithms. Balancing the need to protect agriculturists' privacy with the desire to leverage their data for valuable insights requires careful consideration of data governance policies and regulatory requirements.
4. **Scalability vs. Cost:** As the platform grows in popularity and user base, scalability becomes increasingly important to ensure that it can handle the growing volume of data and user interactions..

CHAPTER 7

CONCLUSION

7.1 CONCLUSION

In conclusion, the development of a web-based platform for agriculturists presents a promising opportunity to revolutionize agricultural practices and empower farmers with valuable insights and decision-making tools. Through the integration of machine learning algorithms, predictive modeling, and decision support mechanisms, the platform offers agriculturists access to timely and accurate information essential for optimizing crop production, pricing strategies, and market timing. However, the success of the platform hinges on effectively addressing various challenges, including limited access to technology, data quality and availability, resource constraints, regulatory compliance, and cultural differences. By carefully considering trade-offs between accuracy and complexity of predictive models, user- friendliness and functionality, data privacy and utility, scalability and cost, and local relevance and generalizability, the development team can design a platform that meets the diverse needs of agriculturists while ensuring usability, scalability, and sustainability.

7.2 RECOMMENDATIONS FOR FUTURE RESEARCH

Moving forward, there are several areas of future research that could further enhance the effectiveness and impact of web-based platforms for agriculturists:

1. **Advanced Predictive Modeling Techniques:** Investigate the application of advanced machine learning techniques, such as deep learning and

ensemble methods, for improving the accuracy and robustness of crop price predictions. Research could focus on developing models capable of capturing complex relationships and patterns in agricultural data, leading to more accurate and reliable forecasts.

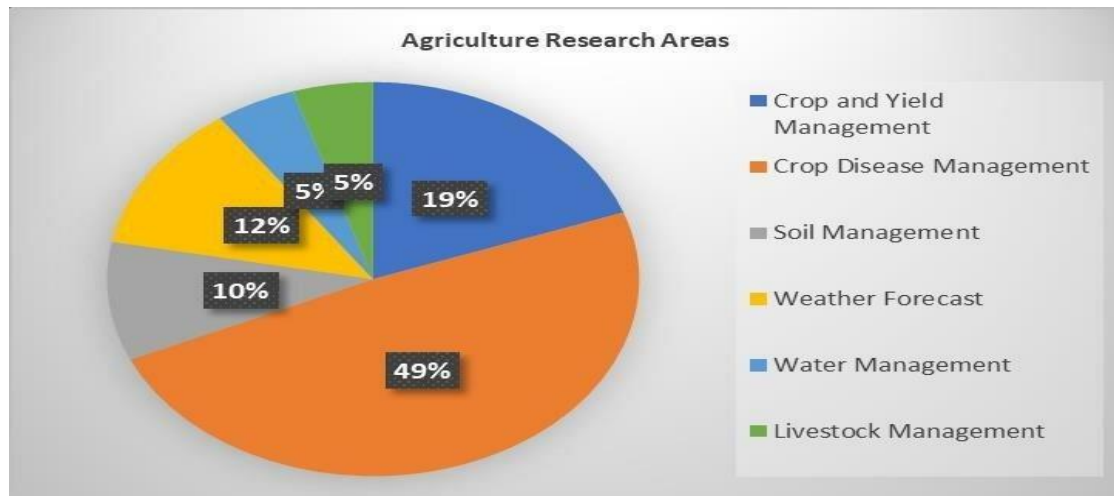


Fig 7.2: Research areas representation

2. Integration of Real-time Data Sources: Explore the integration of real-time data sources, such as satellite imagery, IoT sensors, and social media feeds, to augment existing datasets and provide agriculturists with up-to-the-minute information on weather conditions, crop health, and market trends. Research in this area could focus on developing algorithms and techniques for processing and analyzing large volumes of heterogeneous data streams in real-time.
3. Personalized Decision Support Systems: Investigate the development of personalized decision support systems that leverage machine learning algorithms to provide tailored recommendations and insights to individual agriculturists based on their specific preferences, goals, and constraints. Research could focus on developing algorithms capable of learning from user interactions and feedback to continuously improve the relevance and effective recommendations.

7.3 RESEARCH ISSUES

Certainly, here are some research issues that could be explored in the context of web-based platforms for agriculturists:

1. **Data Quality and Availability:** Investigate methods for improving the quality and availability of agricultural data, including historical crop prices, weather data, soil conditions, and market trends. Research could focus on addressing issues such as data gaps, inaccuracies, and inconsistencies, as well as exploring alternative data sources and data collection methods to augment existing datasets.
2. **Algorithm Selection and Optimization:** Explore the selection and optimization of machine learning algorithms for predictive modeling and decision support in agricultural contexts. Research could focus on comparing the performance of different algorithms, evaluating their suitability for specific tasks and datasets, and developing techniques for optimizing algorithm parameters and hyperparameters.
3. **Interpretability and Explainability:** Investigate methods for enhancing the interpretability and explainability of machine learning models used in web-based platforms for agriculturists. Research could focus on developing techniques for visualizing model predictions, identifying important features and factors driving predictions, and explaining model decisions in a transparent and understandable manner to end- users.
4. **User Interface Design and Usability:** Explore strategies for designing user-friendly interfaces and enhancing usability in web-based

platforms for agriculturists. Research could focus on understanding user needs and preferences, incorporating user feedback into interface design, and developing intuitive and accessible interfaces that facilitate ease of navigation and interaction.

5. Privacy and Security: Investigate approaches for ensuring data privacy and security in web-based platforms for agriculturists. Research could focus on security measures to protect against unauthorized access, data breaches, and cyber-attacks.

REFERENCES:-

- [1] Umamaheswari S, Sreeram S, Kritika N, Prasanth DJ, “BloT: Blockchain-based IoT for Agriculture”, 11th International Conference on Advanced Computing (ICoAC), 2019 Dec 18 (pp. 324-327). IEEE.

- [2] Jain A. “Analysis of growth and instability in the area, production, yield, and price of rice in India”, Journal of Social Change and Development, 2018;2:46-66

- [3] Manjula E, Djodiltachoumy S, “A model for prediction of crop yield” International Journal of Computational Intelligence and Informatics, 2017 Mar;6(4):2349-6363.

- [4] Sagar BM, Cauvery NK., “Agriculture Data Analytics in Crop Yield Estimation: A Critical Review”, Indonesian Journal of Electrical Engineering and Computer Science, 2018 Dec;12(3):1087-93.

- [5] Wolfert S, Ge L, Verdouw C, Bogaardt MJ, “Big data in smart farming—a review. Agricultural Systems”, 2017 May 1;153:69-80.

- [6] Jones JW, Antle JM, Basso B, Boote KJ, Conant RT, Foster I, Godfray HC, Herrero M, Howitt RE, Janssen S, Keating BA, “Toward a new generation of agricultural system data, models, and knowledge products: State of agricultural systems science. Agricultural systems”, 2017 Jul 1;155:269-88.

- [7] Johnson LK, Bloom JD, Dunning RD, Gunter CC, Boyette MD, Creamer NG, “Farmer harvest decisions and vegetable loss in primary production. Agricultural Systems”, 2019 Nov 1;176:102672.

- [8] Kumar R, Singh MP, Kumar P, Singh JP, "Crop Selection Method to maximize crop yield rate using a machine learning technique", International conference on smart technologies and management for computing, communication, controls, energy, and materials (ICSTM), 2015 May 6 (pp. 138-145). IEEE.
- [9] Sriram Rakshith.K, Dr.Deepak.G, Rajesh M, Sudharshan K S, Vasanth S, Harish Kumar N, "A Survey on Crop Prediction using Machine Learning Approach", In International Journal for Research in Applied Science & Engineering Technology (IJRASET), April 2019, pp(3231-3234)
- [10] Veenadhari S, Misra B, Singh CD, "Machine learning approach for forecasting crop yield based on climatic parameters", In 2014 International Conference on Computer Communication and Informatics, 2014 Jan 3 (pp. 1-5). IEEE.
- [11] Ghadge R, Kulkarni J, More P, Nene S, Priya RL, "Prediction of crop yield using machine learning", Int. Res. J. Eng. Technol. (IRJET), 2018 Feb;5.
- [12] Priya P, Muthaiah U, Balamurugan M, "Predicting yield of the crop using machine learning algorithm", International Journal of Engineering Sciences & Research Technology, 2018 Apr;7(1):1-7.
- [13] S. Pavani, Augusta Sophy Beulet P., "Heuristic Prediction of Crop Yield Using Machine Learning Technique", International Journal of Engineering and Advanced Technology (IJEAT), December 2019, pp (135-138)

[14] <https://web.dev/progressive-web-apps/>

[15] <https://www.wikipedia.org/>

[16] Plewis I, "Analyzing Indian farmer suicide rates", Proceedings of the National Academy of Sciences, 2018 Jan 9;115(2): E117.

[17] Nishant, Potnuru Sai, Pinapa Sai Venkat, Bollu Lakshmi Avinash, and B. Jabber. "Crop Yield Prediction based on Indian Agriculture using Machine Learning." In 2020 International Conference for Emerging Technology (INCET), pp. 1-4. IEEE, 2020.

[18] Kale, Shivani S., and Preeti S. Patil. "A Machine Learning Approach to Predict Crop Yield and Success Rate." In 2019 IEEE Pune Section International Conference (PuneCon), pp. 1-5. IEEE, 2019.

[19] Kumar, Y. Jeevan Nagendra, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi. "Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector." In 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 736-741. IEEE, 2020.

[20] Nigam, Aruvansh, Saksham Garg, Archit Agrawal, and Parul Agrawal. "Crop yield prediction using machine learning algorithms." In 2019 Fifth International Conference on Image Information Processing (ICIIP), pp. 125-130. IEEE, 2019.

APPENDIX

A. SOURCE CODE

```
from flask import Flask, render_template, flash, redirect
import joblib
from flask import request
import numpy as np
import pickle
import pandas as pd
from flask import Flask, render_template, request, redirect, url_for
from weather import Weather, WeatherException

app = Flask(__name__)

app.config.from_pyfile('config/config.cfg')
w = Weather(app.config)

model = pickle.load(open('yield.pkl', 'rb'))
crop = pickle.load(open('crop.pkl', 'rb'))

@app.route('/')

@app.route('/index')
def index():
    return render_template('index.html')
@app.route('/login')
def login():
    return render_template('login.html')
@app.route('/chart')
def chart():
    return render_template('chart.html')

@app.route('/upload')
def upload():
    return render_template('upload.html')
@app.route('/preview', methods=["POST"])
def preview():
    if request.method == 'POST':
        dataset = request.files['datasetfile']
        df = pd.read_csv(dataset, encoding = 'unicode_escape')
        df.set_index('Id', inplace=True)
        return render_template("preview.html", df_view = df)

@app.route('/yield_prediction')
def yield_prediction():
    return render_template('yield_prediction.html')

@app.route("/predict", methods = ["POST"])
def predict():
```

```

if request.method == "POST":
    print(request.form)
    State_Name = request.form['State_Name']
    Crop = request.form['Crop']
    Area = request.form['Area']
    Soil_type = request.form['soil_type']

    pred_args = [State_Name,Crop,Area,Soil_type]
    pred_args_arr = np.array(pred_args)
    pred_args_arr = pred_args_arr.reshape(1,-1)
    output = model.predict(pred_args_arr)
    print(output)
    pred=format(int(output[0]))
    Yield= int(pred) / float(Area)
    yields= Yield*1000

    return render_template("yield_prediction.html",prediction_text=pred,
yield_predictions= int(yields))
@app.route('/crop_prediction')
def crop_prediction():
    return render_template('crop_prediction.html')

@app.route('/sandy',methods=['POST'])

def sandy():

    int_features = [float(x) for x in request.form.values()]
    final_features = [np.array(int_features)]

    prediction = crop.predict(final_features)

    preds=format((prediction[0]))

    return render_template("crop_prediction.html",prediction_texts=preds)

@app.route('/homepage')
def homepage():
    return render_template('homepage.html')

@app.route('/result', methods=['POST', 'GET'])
def result_page():
    if request.method == 'POST':
        location = request.form
        w.set_location(location.get('location'))

    try:
        return render_template('result.html', data=w.get_forecast_data())
    except WeatherException:
        app.log_exception(WeatherException)

```

```

        return render_template('error.html')
    else:
        return redirect(url_for('homepage'))
if __name__ == "__main__":
app.run(debug=True)
def page_not_found(error):
    return render_template('404.html'), 404

@app.errorhandler(500)
def internal_server_error(error):
    return render_template('500.html'), 500

# Update predict route with form validation
@app.route("/predict", methods=["POST"])
def predict():
    form = YieldPredictionForm(request.form)
    if request.method == "POST" and form.validate():
        state_name = form.state_name.data
        crop = form.crop.data
        area = form.area.data
        soil_type = form.soil_type.data

        pred_args = [state_name, crop, area, soil_type]
        pred_args_arr = np.array(pred_args).reshape(1, -1)
        output = model.predict(pred_args_arr)
        pred = int(output[0])
        yield_prediction = pred / area * 1000 # Assuming yield in kg per acre

        return render_template("yield_prediction.html", prediction_text=pred,
yield_predictions=int(yield_prediction))
        flash('Please fill in all the fields correctly.')
        return redirect(url_for('yield_prediction'))

# Add CSRF protection
from flask_wtf.csrf import CSRFProtect
csrf = CSRFProtect(app)

if __name__ == "__main__":
app.run(debug=True)

csrf = CSRFProtect(app)

# Configure SQLAlchemy
app.config['SQLALCHEMY_DATABASE_URI'] = 'sqlite:///site.db'
db = SQLAlchemy(app)

# Define models for database
class User(db.Model):
    id = db.Column(db.Integer, primary_key=True)
    username = db.Column(db.String(20), unique=True, nullable=False)
    email = db.Column(db.String(120), unique=True, nullable=False)
    password = db.Column(db.String(60), nullable=False)

```

```

def __repr__(self):
    return f'User('{self.username}', '{self.email}')"

# Create database tables
db.create_all()

# Define a form for crop prediction
class CropPredictionForm(FlaskForm):
    area = FloatField('Area', validators=[DataRequired(), NumberRange(min=0.1)]) # Assuming minimum area is 0.1 acre
    submit = SubmitField('Predict Crop')

# Route for user registration
@app.route("/register", methods=["GET", "POST"])
def register():
    form = RegistrationForm()
    if form.validate_on_submit():
        hashed_password =
bcrypt.generate_password_hash(form.password.data).decode('utf-8')
        user = User(username=form.username.data, email=form.email.data,
password=hashed_password)
        db.session.add(user)
        db.session.commit()
        flash('Your account has been created! You are now able to log in.', 'success')
        return redirect(url_for('login'))
    return render_template('register.html', title='Register', form=form)

# Route for user login
@app.route("/login", methods=["GET", "POST"])
def login():
    form = LoginForm()
    if form.validate_on_submit():
        user = User.query.filter_by(email=form.email.data).first()
        if user and bcrypt.check_password_hash(user.password, form.password.data):
            login_user(user, remember=form.remember.data)
            next_page = request.args.get('next')
            return redirect(next_page) if next_page else redirect(url_for('index'))
        else:
            flash('Login Unsuccessful. Please check email and password', 'danger')
    return render_template('login.html', title='Login', form=form)

# Route for user logout
@app.route("/logout")
def logout():
    logout_user()
    return redirect(url_for('index'))

# Add database integration to yield prediction
@app.route("/predict", methods=["POST"])
@login_required
def predict():
    form = YieldPredictionForm(request.form)
    if request.method == "POST" and form.validate():
        state_name = form.state_name.data

```

```

crop = form.crop.data
area = form.area.data
soil_type = form.soil_type.data

pred_args = [state_name, crop, area, soil_type]
pred_args_arr = np.array(pred_args).reshape(1, -1)
output = model.predict(pred_args_arr)
pred = int(output[0])
yield_prediction = pred / area * 1000 # Assuming yield in kg per acre

# Save prediction to database
prediction = Prediction(state_name=state_name, crop=crop, area=area,
yield_prediction=yield_prediction)
db.session.add(prediction)
db.session.commit()

return render_template("yield_prediction.html", prediction_text=pred,
yield_predictions=int(yield_prediction))
flash('Please fill in all the fields correctly.')
return redirect(url_for('yield_prediction'))

# Add a route for displaying user predictions
@app.route("/predictions")
@login_required
def predictions():
    predictions = Prediction.query.all()
    return render_template("predictions.html", predictions=predictions)

# Add route for deleting a prediction
@app.route("/delete_prediction/<int:prediction_id>", methods=['POST'])
@login_required
def delete_prediction(prediction_id):
    prediction = Prediction.query.get_or_404(prediction_id)
    db.session.delete(prediction)
    db.session.commit()
    flash('Prediction has been deleted!', 'success')
    return redirect(url_for('predictions'))

# Add logging
import logging
from logging.handlers import RotatingFileHandler

if not os.path.exists('logs'):
    os.mkdir('logs')

file_handler = RotatingFileHandler('logs/yourapp.log', maxBytes=10240, backupCount=10)
file_handler.setFormatter(logging.Formatter('%(asctime)s %(levelname)s: %(message)s [in
%(pathname)s:%(lineno)d]'))
file_handler.setLevel(logging.INFO)

app.logger.addHandler(file_handler)
app.logger.setLevel(logging.INFO)
app.logger.info('YourApp startup')

```

```

# Add styling with Bootstrap
# Include Bootstrap CDN links in your HTML templates

# Add error handling for CSRF token
@app.errorhandler(400)
def handle_bad_request(e):
    return render_template('400.html'), 400

# Add a custom 400 error page
@app.route('/400')
def custom_bad_request():
    return render_template('400.html'), 400

# Add route for about page
@app.route("/about")
def about():
    return render_template("about.html", title='About')

# Add route for contact page
@app.route("/contact")
def contact():
    return render_template("contact.html", title='Contact')

    return render_template("privacy.html", title='Privacy Policy')

# Add route for terms of service page
@app.route("/terms")
def terms():
# Add a route for handling 404 errors
@app.errorhandler(404)
def not_found_error(error):
    return render_template('404.html'), 404

# Add a route for handling 500 errors
@app.errorhandler(500)
def internal_error(error):
    db.session.rollback()
    return render_template('500.html'), 500

# Add route for maintenance page
@app.route("/maintenance")
def maintenance():
    return render_template("maintenance.html"), 503

# Add route for 503 errors
@app.errorhandler(503)
def service_unavailable(error):
    return render_template('maintenance.html'), 503

if __name__ == "__main__":
    app.run(debug=True)

```

B.SCREENSHOTS

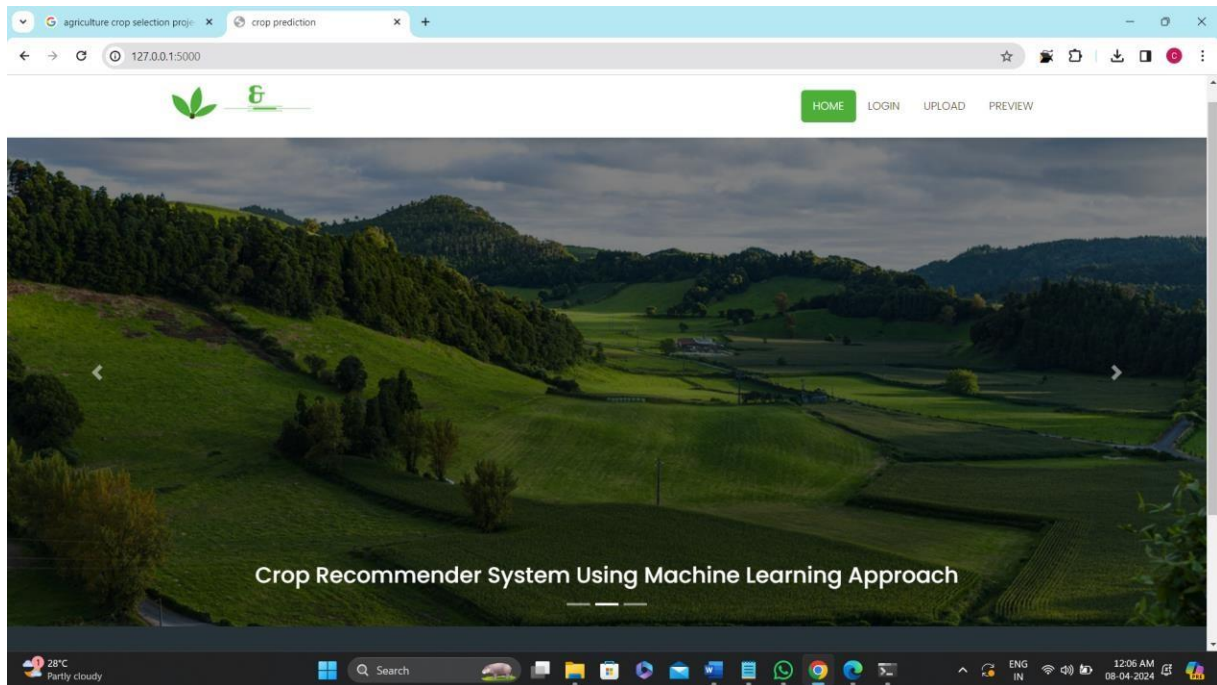


Fig:B.1: Home page

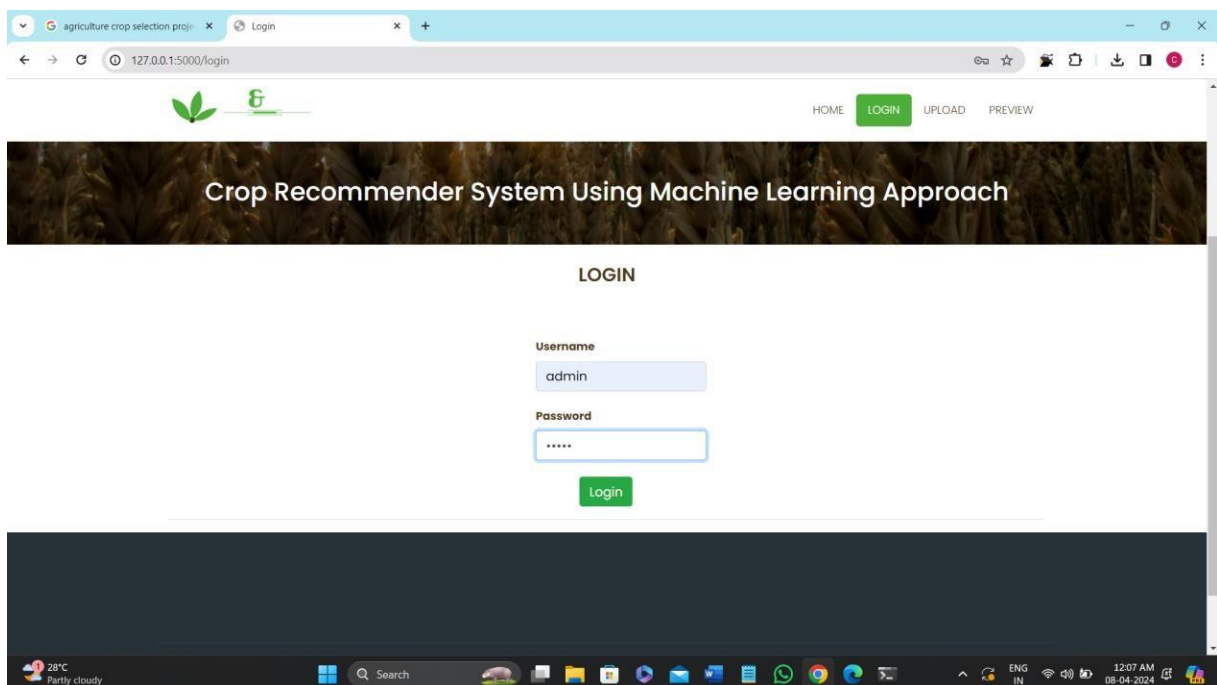


Fig B.2: Login page

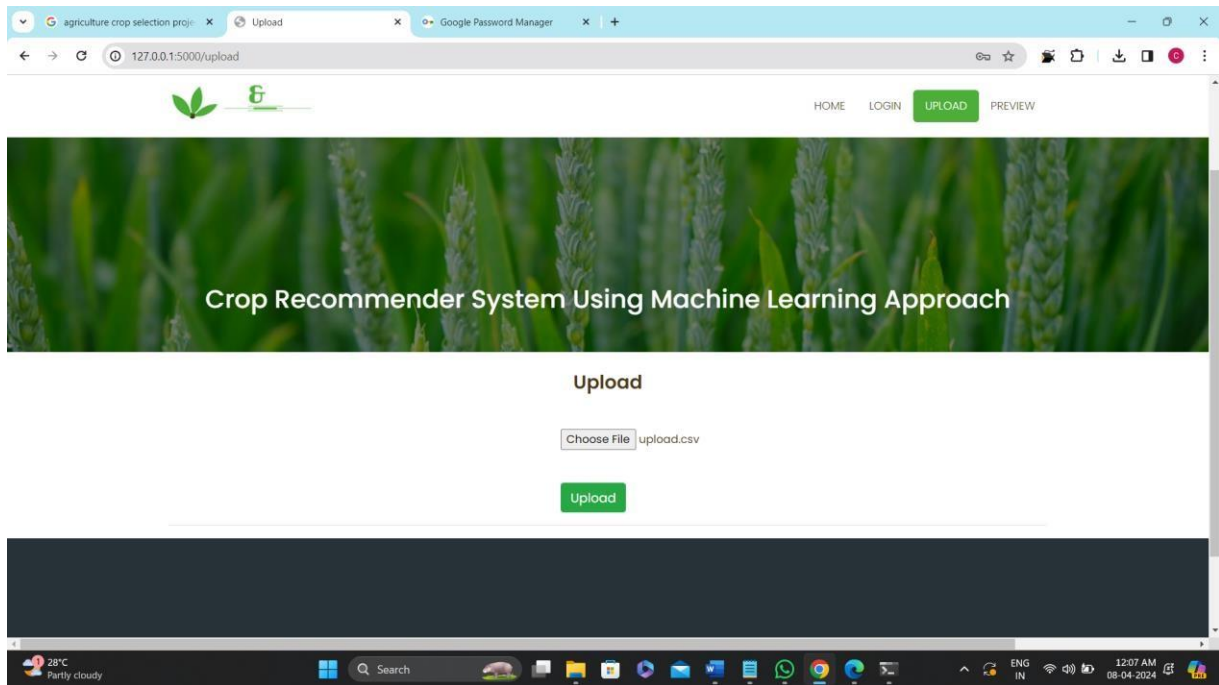


Fig B.3: Upload data

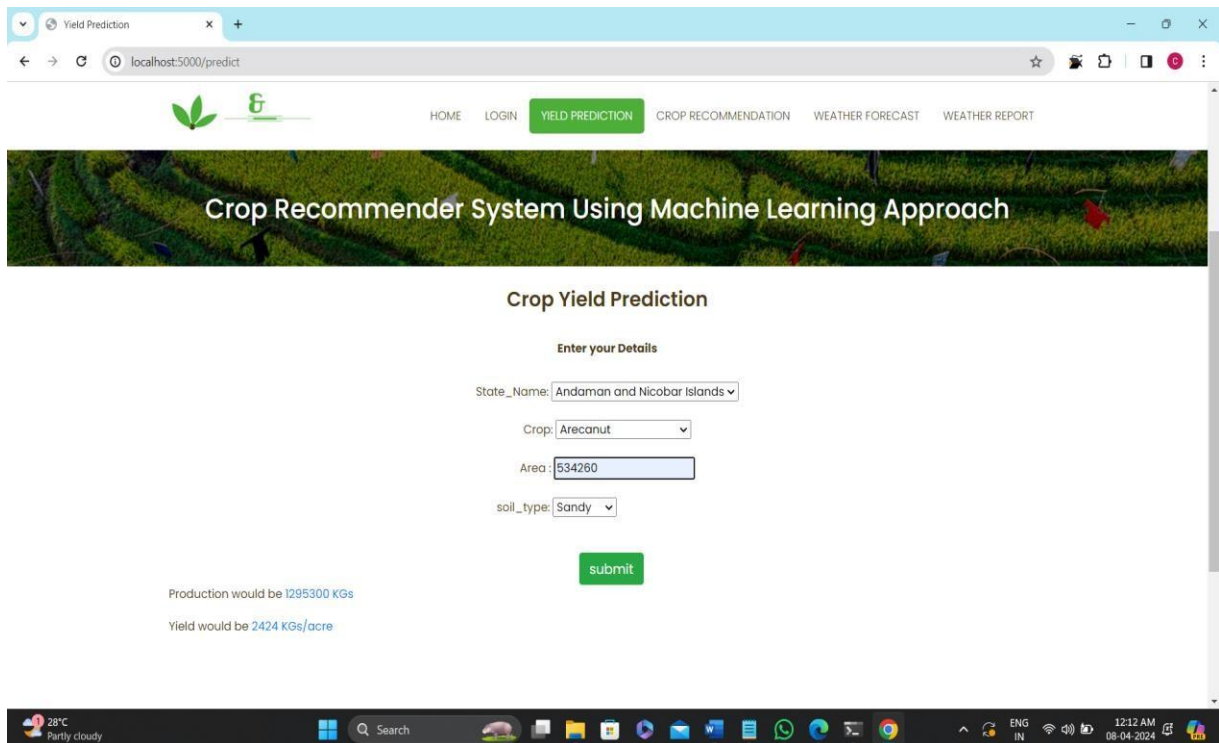


Fig B.4: Crop yield prediction

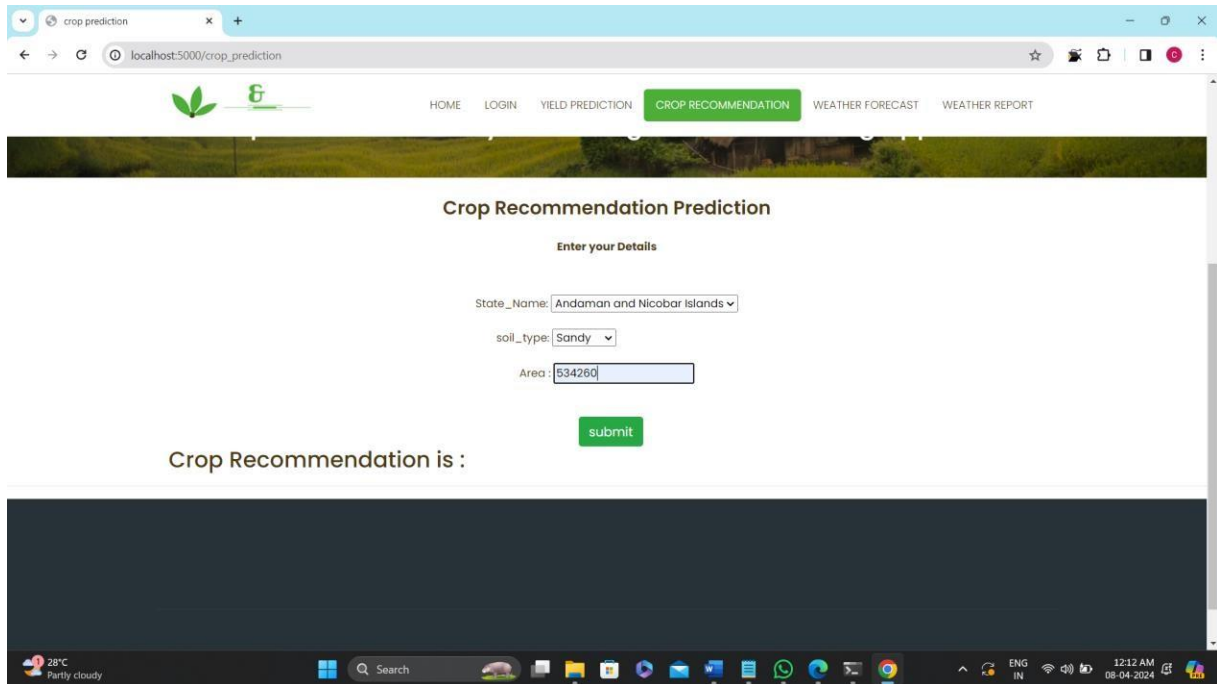


Fig B.5 : Crop Recommendation

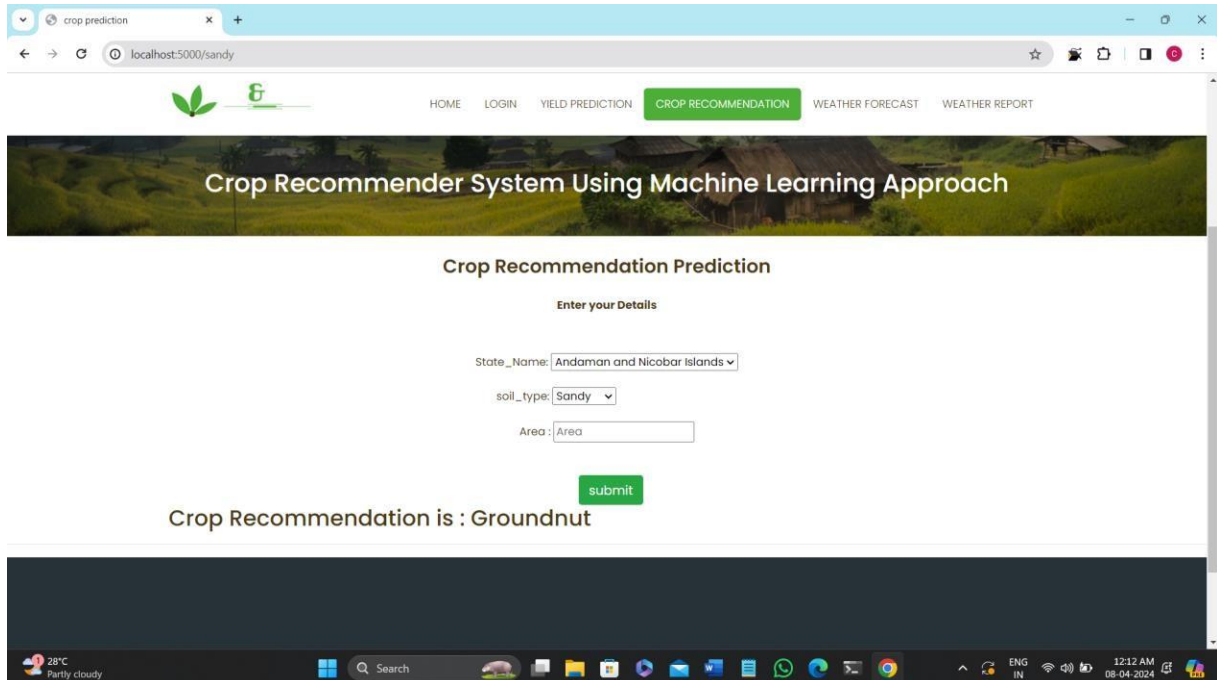


Fig B.6 : Crops based on inputs

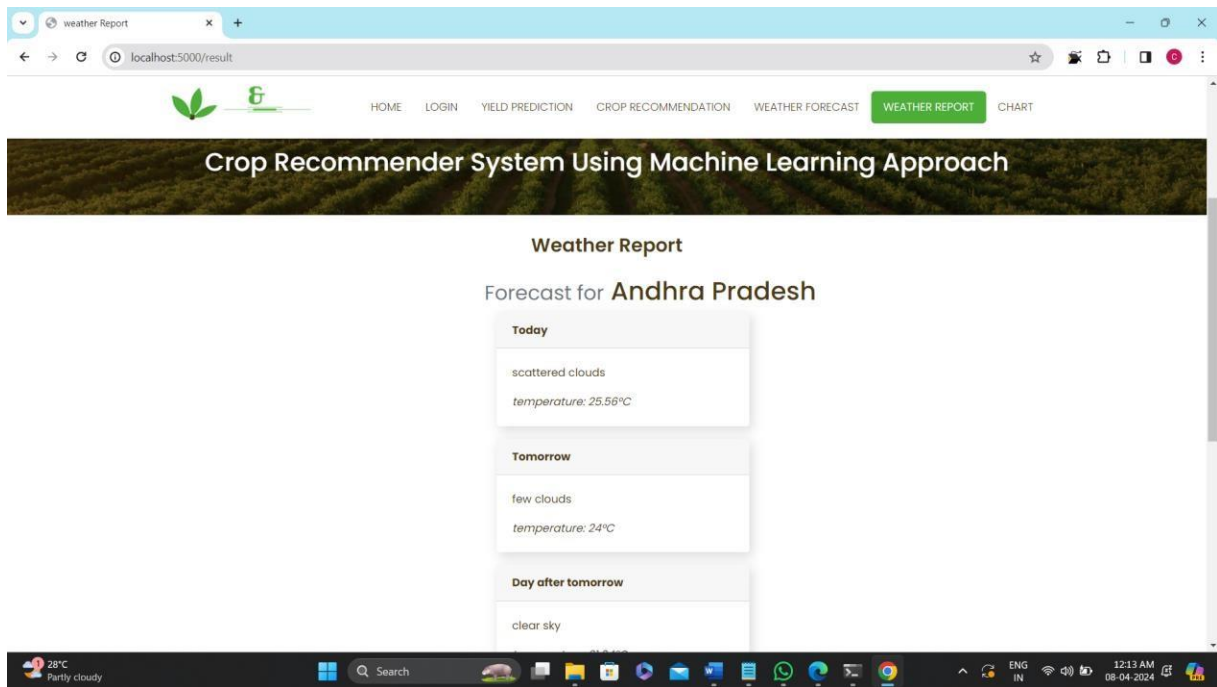


Fig B.7 : Weather Report

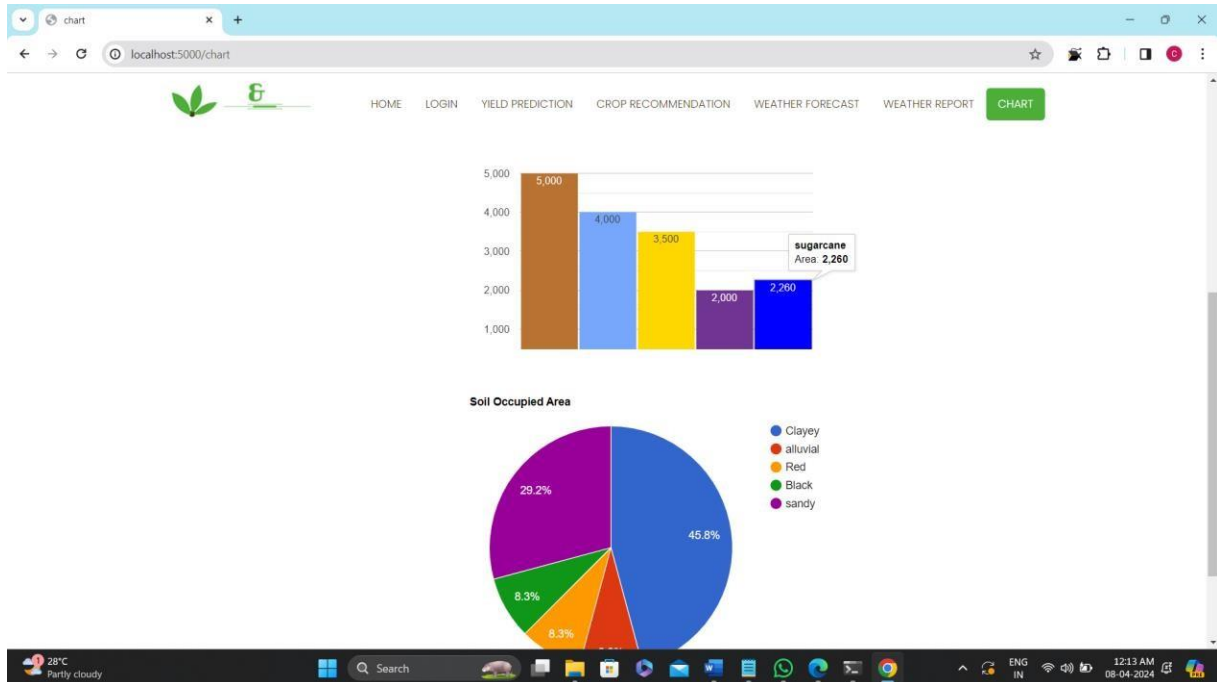


Fig B.8 : Chart Representation

Machine learning-based optimal crop selection system in smart agriculture

CH.Tharun G S Venkata Sai
B.E, Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India.
tharunchilukuri02@gmail.com

D.Sai Jayendra
B.E, Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India
jayendradasari8@gmail.com

A. Mary Posonia

Associate Professor, Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India.
maryposonia.cse@sathyabama.ac.in

ABSTRACT- *In country India, cultivation and related undertakings are certainly the fundamental carriers of life. Additionally, the horticulture industry contributes significantly to the US gross domestic product. The huge size of the rural area is the obligation of the us of a. However, the harvest yield as per hectare is frustrating in contrast with worldwide prerequisites. This is one of the practical explanations behind the exorbitant cost of insolvency among India's minor ranchers. This paper proposes a feasible and rancher cordial yield contraption. The proposed gadget interfaces ranchers through mobile software. GPS enables to find the consumer. Specifies the consumer's region and soil kind as enter. Machine-dominating calculations let you select the most valuable harvest list or are expecting the yield of a harvest chosen through the individual. Selected gadget mastering algorithms like help vector machine (SVM), counterfeit brain local area (ANN), arbitrary backwoods (RF), multivariate straight relapse (MLR) and k-closest (KNN) are utilized to foresee yields.. It. Among them, Irregular Woodland achieved the incredible with 95% exactness. Besides, the machine also proposes the acceptable chance to utilize composts to growth crops.*

Keywords: *Decision timber, random forests, crop choice, yield forecasting.*

I. INTRODUCTION

In a world of latest technologies, fulfilment in statistics sharing will help farmers realise and growth their ability. Information sharing approach that valuable and well-timed information is disseminated to farmers either formally or informally. Willingness to share facts indicates transparency amongst farmers. This openness determines the quantity and quantity of data exchanged. Using internet technologies along with HTML and CSS, we broaden an internet application, accumulate facts from a couple of assets, create statistics structures and put them

where it could expect crop fees. Compilation and type of agricultural products is furnished. They input data into our utility and proportion this facts with farmers, where the records is collected and saved on MySQL servers. Software to ship new records to farmers through SMS immediately. In this manner, farmers will now not should journey to the nearest cities for brand new records. We use machine gaining knowledge of algorithms to expect. Crop prices for the following two months.

For forecasting functions, we will use Vector Machine (SVM), Naive Bayes (NB) and K-Closest (KNN) calculations to foresee the expense of the manufacturing of merchandise. The order system should also be used to make decisions approximately selecting the effects of the classifier.

II. RELATED WORK

[1] Uma keshwari S Farmers can get actual-time facts about seed first-class, weather, answer, soil moisture, and call for and promote charge in a single platform. The reason of this plan is to replace the sensor records in the block. He directed the creation of a clever contract on the ethereal blockchain to make the buy and arrangement of land and yields more straightforward. With authorization to steady information on seed quality, weather conditions, and soil conditions, ranchers can reach extra educated final products about crop choice, spreading out plans, water system, and course of action. Crop yields are extended and resource dissemination is further developed consequently. By seeing parts, for example, ecological conditions and soil moisture levels consistently, farmers can better anticipate and mitigate risks such as crop diseases, pest infestations, and adverse weather events. Early detection and response can help minimize crop losses and protect farmers' livelihoods.

The integration of market demand and pricing data into the blockchain platform enables farmers to make informed decisions about identify market trends, assess demand-supply dynamics when and where to sell their produce. [2] P. Parameswari- The outcomes show that the PART calculation performs well, with the most elevated accuracy of the crop is 98.33% when contrasted with different techniques, and it consumes lesser chance to assemble the model. [3] Sandhya Umrao - This study aims to show that the machine learning method is most suitable for predicting agricultural production in a certain area, state, or season. [4] R. Kavitha Accuracy horticulture has given an answer for this issue for ranchers. Agribusiness has become huger because of accuracy farming. [5] M Aruna Devi - The result of Irregular Woodland is contrasted and Backing Vector Machine calculation the creators presumed that the proposed model deals with normal exactness of 90%.

III. PROPOSED WORK

In this proposition, we suggest a PC based absolutely rendition issues. The new gadget is intended to help ranchers Increment efficiency and afterward offer an extra productive harvest of a chose locale. The proposed model presents crop decision on monetary and financial premise Ecological circumstances and gifts to development crop yield. Then assist with meeting the creating call for Supplements the proposed form predicts crop yield through dissecting Factors like precipitation, temperature, region, season, soil type. The machine permits decide the best chance to utilize manure. Client manages the cost of soil area to be developed and enter soil type. The proposed adaptation predicts crop yield the exceptional love form additionally suggests the most valuable vegetation It lets you know the appropriate opportunity to apply manure. The principal objective is to development the sort of harvests that can be made. The proposed framework will help control more prominent time Challenges defied by involving the rancher in decision of vegetation and most extreme yield.

There are six categorical columns in the final data frame: State, District, Season, Crop, and Soil Type. As a result, a dataset can be divided into two types of variables: continuous variables and categorical variables. A categorical variable takes one value from a limited set of values and is not quantifiable.

SYSTEM ARCHITECTURE:

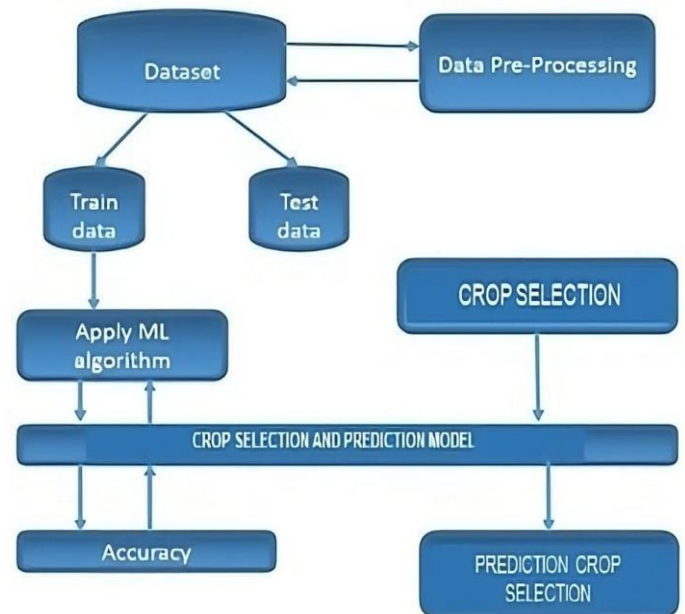


Fig 1: System Architecture

ALGORITHM:

1) Initialization:

- 1.1) Import required libraries:
 - Pandas
 - Scikit-learn
 - Matplotlib
 - Seaborn
- 1.2) Initialize datasets:
 - Historical crop data
 - Weather patterns
 - Soil characteristics
 - Previous yield records
- 1.3) Initialize the Random Forest model.

2) Data Preprocessing:

- 2.1) Merge and clean datasets.
- 2.2) Perform feature engineering:
 - Normalize numerical features
 - Encode categorical variables
 - Handle missing values

If we have a numerical feature X , we can normalize it using min-max scaling:

$$X_{\text{normalized}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}}$$

3) Model Training:

- 3.1) Split the dataset into training and testing sets.
 - 3.2) Train the Random Forest model on the training data.
 - 3.3) Tune hyperparameters using cross-validation.
- Calculation of temperature and rainfall during training of the model:
- $$\eta_1(T) = \exp \{ -a ((T - T_{opt}) / 10) \}$$

4) Crop Selection:

- 4.1) Input new data for prediction.
- 4.2) Utilize the trained model to predict suitable crops.
- 4.3) Output recommended crops based on results.

5) Yield Forecasting:

- 5.1) Collect current environmental data (e.g., soil).
- 5.2) Input environmental data into the trained model.
- 5.3) Predict crop yields for the upcoming season.
- 5.4) Generate yield forecasts and confidence intervals.

6) Evaluation:

- 6.1) Assess the accuracy of crop selection and yield forecasting.
- 6.2) Compare model predictions with actual outcomes.
- 6.3) Analyze any discrepancies and potential improvements.

The F_{TOA} flux is calculated depending on the season and the zenith angle of the Sun:

$$F_{TOA}(t_d) = F_0 \left[1 + 0.033 \cos \left(\frac{360 t_d}{365} \right) \right] \langle \cos \theta_s \rangle$$

7) Visualization:

- 7.1) Plot feature importance to understand key factors influencing predictions.
- 7.2) Visualize yield forecasts over time.
- 7.3) Create interactive dashboards for stakeholders.

Growing season is carried out on the basis of the following Equation:

$$Y = \gamma \sum_{i=1}^N F_{TOA} \times NDVI_i \times \eta_1(T_i) \times \eta_2(W_i) \times (1 \cdot (T_i))$$

One of the key advantages of Random Forest is its ability to provide insights into feature importance. By Analyzing how much each feature contributes to the reduction in impurity across all trees, the algorithm assigns importance scores to features. Features that lead to a greater reduction in impurity are considered more important in making predictions.

This feature importance analysis aids in understanding the underlying patterns in the data and selecting relevant features for model interpretation and decision-making. Overall, Random Forest is a versatile and powerful algorithm widely used for classification, regression, and feature selection tasks in machine learning.

MODULES DESCRIPTION:

3.1 User interface and data collection

This is the first activity, User needs to provide a correct contact number and a password, which user enters while registering, in order to login into the webpage. If information provided by the user matches with the data in the database table then user successfully login into the webpage else message of login failed is displayed and user need to reenter correct information. A link to the register activity is also provided for registration of new users.

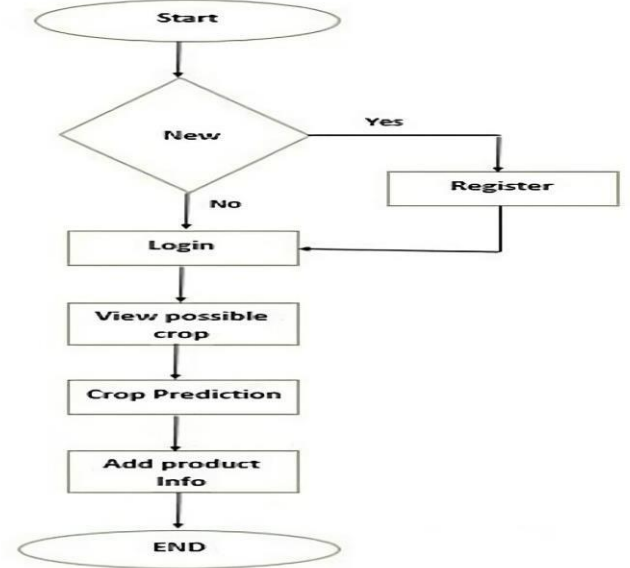


Fig 4: user interface design

3.2 Data Preprocessing

Hear the raw data in the crop data is cleaned and the metadata is appending to it by removing the things which are converted to the integer. So, the data is easy to train. Hear all the data. In this preprocessing, we first load the metadata into this and then this metadata will be attached to the data and replace the converted data with metadata .



Fig 5: Processing the collected data

3.3 Prediction and Analysis

The obtained result will be helpful for the farmers to know the Yield of the crop so, he can go for the better crop which gives high yield and also say them the efficient use of agriculture field. This way we can help the farmers to grow the crop which gives them better yield.

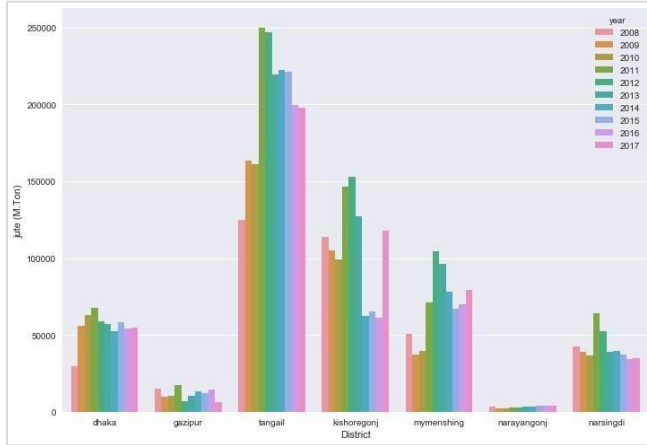


Figure 6. Production of Jute rice

Average rainfall is a crucial factor in agriculture crop selection and yield forecasting as it directly impacts the growth and productivity of crops. Farmers and agricultural experts often consider historical rainfall data to make informed decisions about which crops to plant and when to expect yields. Additionally, understanding rainfall patterns enables farmers to implement appropriate irrigation strategies and select drought-resistant crop varieties to mitigate the effects of water scarcity. Accurate rainfall forecasting also aids in optimizing resource allocation and maximizing agricultural productivity while minimizing risks associated with adverse weather conditions.

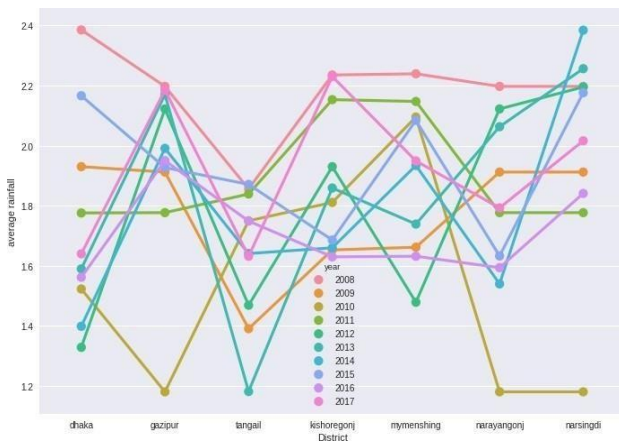


Fig 7 : Average Rainfall

IV. RESULT AND DISCUSSIONS

Between SVM and Decision Tree Classifier algorithms, SVM gave the very last accuracy of forty four.67%, whilst here it's miles 86.Ninety two. The SVM did not improve as the information set improved, but the version endured to carry out worse. Since there is lots of noise inside the information (i.e. the target lessons overlap) [1] there can be situations wherein 2 or extra inputs are similar, the SVM does now not work as anticipated. Random Forest Regressor predicts returns with ninety two.02% accuracy.

states	rainfall	Temperature	Soil type	crop
Andhra pradesh	116.75	31.5	Clayey	Paddy-19795
Tamil nadu	110.98	32.6	Alluvial	Sugarcane-197528
Arunachal Pradesh	325.50	32.38	Clayey	Ragi-204856
Punjab	425.38	29.5	Sandy Loam	Rice-185659

Table 1 : attributes of crops

Crop yield prediction is a complex process which relies on several different factors including weather, soil, fertilizers, pest infestations, etc. In this paper, we predict the crop yield using weather and soil parameters. The research is based on the datasets limited to districts in India. The system incorporates regression techniques to estimate yield and multi-class classification to predict type of the crop. Also, various factors like irrigation and fertilizers use can be included for better prediction. Mobile App can be developed for mobile devices with added services like price estimates in accordance with current market prices. Paid datasets may bring more reliable and accurate data which in turn might help in model accuracy. They may contain more features which may help correlate more with label.

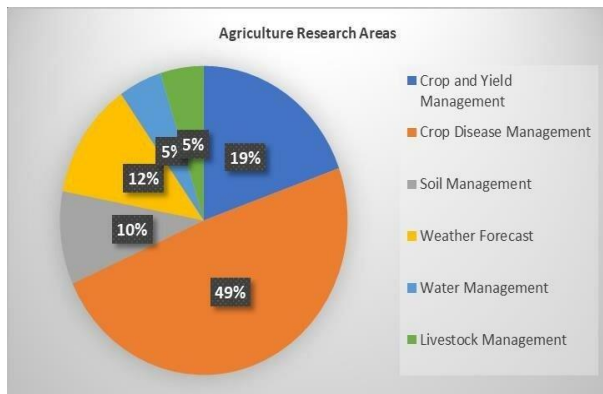


Fig 8: Crop Management

Effective crop management encompasses various practices such as soil fertility management, pest and disease control, and proper crop rotation. Integrating crop selection and yield forecasting into management strategies allows farmers to make informed decisions regarding planting schedules, input application, and harvest planning. By utilizing data-driven approaches, farmers can optimize resource utilization, enhance crop health, and ultimately improve yields for sustainable agricultural production.

V. CONCLUSION

This research paper goes into great length about the significance of coping. These days, they aid farmers in maximizing their output. Farmers are able to supply timely vegetation. Numerous instrumental methods have been employed to examine the rural environment. The literature review looks at a few tactics related to different agricultural problems. The development of new neural networks and soft computing methods is crucial to the creation of guidelines. Farmers may receive more individualized and pertinent advice for raising yields by considering characteristics in addition to productivity and environment.

Furthermore, the integration of coping mechanisms with advanced technologies offers promising opportunities for agricultural innovation. Precision agriculture, for instance, leverages data analytics, remote sensing, and automation to

optimize resource utilization and decision-making processes on the farm. By harnessing the power of these technologies, farmers can implement precise interventions tailored to specific crop needs and environmental conditions. This not only enhances productivity but also promotes sustainability by minimizing inputs and reducing environmental impact. The collaborative efforts of researchers, policymakers, and industry stakeholders are essential for fostering the adoption of such innovative approaches and ensuring their scalability and accessibility to farmers worldwide.

V. REFERENCES

- [1]. Mary P.M, and V. L. Jyothi. "XML document retrieval by developing an effective indexing technique." In 2014 Sixth International Conference on Advanced Computing (ICoAC), pp. 120-123. IEEE, 2014.
- [2]. Keerthana, K. T. E., S. Karpagavalli, and A.M Psonia. "Smart system monitoring agricultural land Using IoT." In 2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research (ICETIETR), pp. 1-7. IEEE, 2018.
- [3]. Akhila, Chitturi, Cheemala Astha Diamond, and A.P Mary. "Convolutional neural network based online rail surface crack detection." In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1602-1606. IEEE, 2021.
- [4]. Lakshmanan, Gowri, and Mary M.P "A novel analysis on application of neural support on nuclear reactor control process monitoring." Indian J. Sci. Technol 9, no. 10 (2016).
- [5]. Psonia, A. Mary, and D. V. Jyothi. "Improving data access performance by reverse indexing." *Int. J. Eng. Technol.(IJET)* 7, no. 3 (2015).
- [6]. Govindarajan, Sreevatsan, Sujeeth Sundarajan Rajkumar, and P.A Mary. "Blockchain Fundraising and Charity Platform." In 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN), pp. 1-6. IEEE, 2023.
- [7]. M.A Psonia, A., Ankayarkanni, B., Usha Nandhini, D., Albert Mayan, J. and Nagarajan, G., 2022. An Efficient Algorithm for Traffic Congestion Control. In *Advances in Intelligent Computing and Communication: Proceedings of ICAC 2021* (pp. 441-448). Singapore: Springer Nature Singapore.
- [8]. Nekkanti, Lakshmi Bhavani, A. Priyanga, A. Mary Psonia, and J. Albert Mayan. "Breaking Down Communication Barriers: Real-Time Sign Language Recognition Using CNN & Flask-Based API." In 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT), pp. 330-336. IEEE, 2023.
- [9]. Mercy.Paul Selvan "XML document retrieval by developing an effective indexing technique." In 2014 Sixth International Conference on Advanced Computing (ICoAC), pp. 120-123. IEEE, 2014.
- [10]. D. Anto Praveena ."XML document retrieval by developing an effective indexing technique." In 2014 Sixth International Conference on Advanced Computing (ICoAC), pp. 120-123. IEEE, 2014.



PANIMALAR ENGINEERING COLLEGE

An Autonomous Institution & Affiliated to Anna University, Chennai
Approved by AICTE, New Delhi, Accredited by National Board of Accreditation(NBA)
Approved by UGC for 2(f) & 12(B) Status
Bangalore Trunk Road, Poonamallee, Chennai-600123.

Certificate
of
Participation



PECTEAM 2K24



This is to certify that Dr./Mr./Ms.....**CH.THARUN.G.S.VENKATA SAI**.....of
.....**CSE**.....from **Sathyabama Institution of Science and technology**...has presented paper
entitled.....**MACHINE LEARNING-BASED OPTIMAL CROP SELECTION SYSTEM IN SMART**.....
AGRICULTURE..... in 7th International Conference on
Intelligent Computing(IConIC) held on 22nd & 23rd MARCH,2024.

FACULTY CO-ORDINATOR

CO - CONVENER

Prof. S. VIMALA, M.TECH., (Ph.D)

CO - PATRON

Dr.K.MANI, M.E., Ph.D.,

