

AGRICULTURE YIELDING USING ML TECHNIQUES

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

by

KONA JAYASARIKA (37110354)



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

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

**Accredited with Grade "A" by NAAC | 12B Status by UGC | Approved by
AICTE**

JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI – 600 119

March – 2021



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited with "A" grade by NAAC | 12B Status by UGC | Approved by
AICTE

Jeppiaar Nagar, Rajiv Gandhi Salai,
Chennai – 600 119

www.sathyabama.ac.in

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of
K. JAYASARIKA (Reg. No 37110354) who carried out the entitled
"AGRICULTURE YIELDING USING ML TECHNIQUES " under my
supervision from **November 2020 to March 2021.**

Internal Guide

Dr. S. Murugan, M.E., Ph.d.,

Head of the Department

Submitted for Viva voce Examination held on _____

Internal Examiner

External Examiner

DECLARATION

I **KONA JAYASARIKA** hereby declare that the Project entitled “**AGRICULTURE YIELDING USING ML TECHNIQUES**” is done by me under the guidance of **Dr. S. Murugan, M.E.,Ph.d.**, Department of Computer Science and Engineering at Sathyabama Institute of Science and Technology is submitted in partial fulfilment of the requirements for the award of Bachelor of Engineering degree in Computer Science and Engineering.

DATE :

**PLACE :
CANDIDATE**

SIGNATURE OF THE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board of Management of SATHYABAMA** 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 and Dr.L.Lakshmanan M.E., Ph.D.,and Dr.S.Vigneshwari M.E., Ph.D., Heads 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 Mr. for his 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

Crop yield of a country has a direct impact on the economy of that country as well as on its overall development. It plays a very important role in determining every aspect that effects the growth of a country i.e., from insufficient yield to the economic trading of crops. So to understand the various factors that affects the crop yield as well as in order to predict the yield of the crop based on these factors we make use of machine learning algorithm and various packages that are available in it. In this approach, we consider the important factors that has the direct effect on the crop yield for analysis and prediction. For example, the basic factors on which the yield of crops depends are rainfall, temperature, humidity, soil characteristics, fertilizers, wind speed, sunshine and also on the farming practices. These factors varies for various crops as well as for various locations. Finally it will recommend the crop based on input values.

TABLE OF CONTENTS

ABSTRACT	v
LIST OF FIGURES	viii

CHAPTER No.	TITLE	PAGE No.
1	INTRODUCTION	1 - 2
	1.1 INTROUCTION	1
2	LITERATURE SURVEY	3 - 6
	2.1 VIRENDRA PANPATIL et	3
	2.2 MAYANK et	3
	2.3 SHWETA et	3
	2.4 AMIT KUMAR et	4
	2.5 MANJULA et	4
	2.6 RAKESH KUMAR et	5
	2.7 RAJSHEKHAR et	5
	2.8 VISHNUVARDHAN et	6
3	AIM AND SCOPE	7 - 9
	3.1 AIM OF THE PROJECT	7
	3.1.1 OBJECTIVES	7
	3.2 SCOPE	7
	3.2.1 ADVANTAGES	7
	3.2.2 DISADVANTAGES	7
	3.3 FEASIBILITY STUDY	8
4	EXPERIMENTAL METHODS AND ALGORITHMS	10 - 17
	4.1 HARDWARE REQUIREMENTS	10
	4.2 SOFTWARE REQUIREMENTS	10
	4.3 PYTHON LANGUAGE	10
	4.3.1 CHARACTERISTICS	11
	4.3.2 APPLICATIONS	11

	4.4 SYSTEM DESIGN	13
	4.4.1 ARCHITECTURE	13
	4.4.2 UML DIAGRAMS	13
	4.4.3 GOALS	14
	4.4.4 USE CASE DIAGRAMS	14
	4.4.5 SEQUENCE DIAGRAMS	15
	4.4.6 ACTIVITY DIAGRAMS	16
5	RESULT AND PERFORMANCE ANALYSIS	18 - 24
	5.1 MODULE DESCRIPTION	18
	5.1.1 DATA PRE- PROCESSING	18
	5.1.2 DATA SET DESCRIPTION	19
	5.1.3 ALGORITHM	19
	5.2 SOFTWARE TESTING	20
	5.2.1 UNIT TESTING	21
	5.2.2 INTEGRATION TESTING	21
	5.2.3 FUNCTIONAL TESTING	21
	5.2.4 SYSTEM TESTING	21
	5.2.5 WHITE BOX TESTING	22
	5.2.6 BLACK BOX TESTING	22
	5.2.7 ACCEPTANCE TESTING	22
	5.3 SCREENSHOTS	23
6	CONCLUSIONS AND FUTURE WORK	25
	6.1 CONCLUSION	25
	6.2 FUTURE WORK	25
	REFERENCES	26 - 27
	APPENDIX	28 - 38
	A. PAPER ACCEPTANCE MAIL	28
	B. PLAGIARISM REPORT	29
	C. JOURNAL PAPER	30
	D. SOURCE CODE	31

LIST OF FIGURES

FIGURE No.	FIGURE NAME	PAGE No.
4.4.1	System Architcture	13
4.4.4	Use Case Diagram	15
4.4.5	Sequence Diagram	16
4.4.6	Activity Diagram	17
5.3.1	Data Entry Page	23
5.3.2	Output for some data values	23
5.3.3	Graph(o/p)	24

CHAPTER – 1

INTRODUCTION

1.1 INTRODUCTION :

Machine Learning:

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves. The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Machine Learning Methods:

Machine learning algorithms are often categorized as supervised or unsupervised.

Supervised machine learning algorithms:

It can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.

Unsupervised machine learning algorithms:

This are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right

output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.

Semi-supervised machine learning algorithms:

This fall somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training – typically a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semi-supervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiring unlabeled data generally doesn't require additional resources.

Reinforcement machine learning algorithms:

It is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behaviour within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal.

CHAPTER – 2

LITERATURE SURVEY

2.1 VIRENDRA PANPATIL ET :

It had accomplished gigantic work for Indian ranchers by making productive yield proposal framework. They created framework utilizing classifier models, for example, Decision Tree Classifier, KNN, and Naive Bayes Classifier. The proposed framework can be utilized to figure out best season of planting, development of plant and Plant reaping. They utilized distinctive classifier for accomplishing better exactness for instance: Decision tree shows less precision when dataset is having more varieties yet Naïve Bayes gives preferable exactness over choice tree for such datasets. The best favorable position of framework that it can without much of a stretch versatile all things considered/be utilized to test on various yields.

2.2 MAYANK ET :

It have presumed that this paper fabricate extemporized framework for crop yield utilizing administered AI calculations and with objective to give simple to utilize User Interface, increment the precision of crop yield forecast, investigate distinctive climatic boundaries, for example, overcast cover, precipitation, temperature, and so on In the proposed framework they zeroed in on MAHARASHTRA State for implantation and for information gathering they utilized govt. site, for example, www.data.gov.in. For crop yield forecast they utilized calculations, for example, Random Forest Algorithm and for convenience they created website page so it will be not difficult to use for all. The primary favourable position of proposed framework is precision rate is more than 75 percent on the whole the yields and areas chose in the examination.

2.3 SHWETA ET :

It have inferred that this paper will survey that different utilizations of AI in the cultivating areas. And furthermore, helps in can be select appropriate crop

select land and select season settled utilizing these procedures. The calculations use are Naive Bayes and K-Nearest Neighbor. The calculations are utilizes precision of execution.

2.4 AMIT KUMAR ET :

It have presumed that this paper helps in foreseeing crop arrangements and augmenting yield rates and making advantages to the ranchers. Additionally, Using Machine learning applications with farming in foreseeing crop sicknesses, examining crop copies, diverse water system designs. The calculations utilized are fake neural organizations. The serious issue with neural organization is that the proper organization which suits best for the arrangement is difficult to accomplish and it incorporates experimentation. The second issue with neural organization is the equipment reliance as the calculation incorporates more calculations in reverse and forward the preparing needs more. Assurance of appropriate organization structure requires insight and time. The proposed framework likewise centers around crop determination utilizing natural just as financial variables. The framework likewise utilizes the monetary factor that is the cost of the crop which assumes a significant part on the off chance that if the yields with same yield yet unique yield cost. The framework additionally utilizes other strategy which is crop sequencing which gives a full arrangement of yield which can be developed all through the season. The proposed framework likewise centers around crop choice utilizing ecological just as financial variables. The framework likewise utilizes the monetary factor that is the cost of the crop which assumes a significant part on the off chance that if the crops with same yield yet unique yield cost. The framework additionally utilizes other technique which is crop sequencing which gives a full arrangement of yield which can be developed all through the season.

2.5 MANJULA ET :

It has have presumed that this paper helps in improving the yield pace of crops by utilizing rule based mining. The paper utilizes affiliation rule mining to foresee the yield of the crop. The calculations utilized are k-Means Algorithm, bunching strategy and deduced affiliation rule mining. The significant impediment is that the paper utilizes affiliation rule digging for expectation of crop yield. The issue

with affiliation decide mining is that it creates an excessive number of rules sometimes and the exactness of the expectation decreases. Likewise the principles will in general fluctuate according to dataset and the outcomes additionally enormously. The proposed framework mostly centers around the issue of yield expectation of crop which assumes vital part in yield choice as rancher can choose crop with greatest yield. The frameworks utilizes affiliation rule mining to discover rules and crops with greatest yield. This framework centers around formation of an expectation model which might be utilized to future forecast of crop yield.

2.6 RAKESH KUMAR ET :

It has have presumed that this paper helps in improving the yield pace of crops by applying order techniques and looking at the boundaries. The paper clarifies the utilization of various calculations to accomplish the equivalent. The calculations proposed are Bayesian calculation, K-implies Algorithm, Clustering Algorithm, Support Vector Machine. The hindrance is that there could be no appropriate precision and execution referenced in the paper according to usage of the proposed calculations. The paper is a study paper and just recommends the utilization of the calculations yet there is no usage proof gave in the paper. The technique applied on this paper for crop decision centers uniquely around the plants which might be developed as indicated by season. The proposed approach settle decision of crop (s) principally dependent on forecast yield cost supported by boundaries (for example Environment, soil kind, water thickness, crop kind). It takes crop, their planting time, estate days and foreseen yield charge for the season as information and finds a succession of vegetation whose creation with regards to day are greatest over season.

2.7 RAJSHEKHAR ET :

The depicts and gave the subtleties us for rundown of utilized techniques, In India there are divergent Agriculture crops creation and those crops relies upon the few sort of elements, for example, natural science, economy and furthermore the geological variables covering such procedures and strategies on memorable yield of disparate yields, it is conceivable to get information or information which can be steady to ranchers and government associations for creation well choices and for improve rules which help to expanded creation. In this article, our work is

on utilization of information mining strategies which is use to separate data from the horticultural records to assess better crop yield for primary yields in principle regions of India. In our task we found that the exact expectation of disparate indicated crop yields across various locale will help to ranchers of India. From this Indian ranchers will plant various crops in various distr [8].

2.8 VISHNUVARDHAN ET :

They examined a few development in India is dealing with thorough issue to benefit as much as possible from the crop efficiency. More than 60 out of a hundred the crop actually relies upon rainstorm precipitation. Momentum developments in Information Technology for Agriculture field have built up an intriguing exploration zone to conjecture the crop yield. The risky of yield expectation is a significant issue that stays to be addressed dependent on available information. Information mining techniques are the better determinations for this reason. Distinctive Data Mining strategies are utilized and assessed in agribusiness for approximating the impending year's crop creation. This paper presents a concise investigation of crop yield forecast utilizing Multiple Linear Regression (MLR) strategy and Density based grouping procedure for the specific district for example East Godavari region of Andhra Pradesh in India. In this paper an exertion is made in order to know the locale exact crop yield examination and it is prepared by applying both Multiple Linear Regression technique and Density-based bunching strategy. These models were tested in regard of the multitude of areas of Andhra Pradesh, at that point the strategy of assessment is dropped with just East Godavari region of Andhra Pradesh in India [1].

CHAPTER – 3

AIM AND SCOPE

3.1 AIM OF THE PROJECT :

For the better crop yield, the farmers always require a timely guidance to predict the future of crop productivity and also an analysis is to be made which will help the farmers to utilize full capacity in the crop production for their crops. Yield prediction is a vital agricultural problem. The volume of data is vast in Indian agriculture. For agriculture problems data mining is applied widely. As every farmer is interested by knowing that how much yield is expected to make. In the past, with the farmers previous experience for a particular crop, one can make the predictions for crop.

3.1.1 Objective :

In this project we are using some module and algorithms to predict the crop yield and for good crop growth.

3.2 SCOPE :

Applying Naive Bayes Data Mining Technique for Crop selection will depending on the nature of the Naive probability model. It can be trained very easy in a supervised learning section. In several practical applications, parameter estimation for naive Bayes uses the method of naive Bayes model with believing in Bayesian probability or using any Bayesian methods.

3.2.1 Advantages :

1. A supervised classifier can perform tasks which linear program cannot.
2. It works even in the presence of noise with good quality output.

3.2.2 Disadvantages :

1. Must have Knowledge on Bayesian probability or Bayesian methods.
2. Time taken for the process is larger.
3. Based on the assumption that features have same statistical relevance.

3.3 FEASIBILITY STUDY :

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

- Economic 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 being placed on the client. A feasibility study evaluates the project's potential for success.

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.

CHAPTER – 4

EXPERIMENTAL METHODS AND ALGORITHMS

4.1 HARDWARE REQUIREMENTS :

The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware. The minimal hardware requirements are as follows,

- | | |
|--------------------|--------------|
| 1. Processor | : Pentium IV |
| 2. RAM | : 8 GB |
| 3. Processor | : 2.4 GHz |
| 4. Main Memory | : 8GB RAM |
| 5. Hard Disk Drive | : 1tb |
| 6. Keyboard | : 104 Keys |

4.2 SOFTWARE REQUIREMENTS :

Software requirements deals with defining resource requirements and prerequisites that needs to be installed on a computer to provide functioning of an application. The minimal software requirements are as follows,

- | | |
|---------------------|--------------|
| 1. Front end | : python |
| 2. Dataset | : csv |
| 3. IDE | : anaconda |
| 4. Operating System | : Windows 10 |

4.3 PYTHON LANGUAGE :

Python is an object-oriented programming language created by Guido Rossum in 1989. It is ideally designed for rapid prototyping of complex applications. It has interfaces to many OS system calls and libraries and is extensible to C or C++. Many large companies use the Python programming language include NASA, Google, YouTube, BitTorrent, etc. Python programming is widely used in Artificial Intelligence, Natural Language Generation, Neural Networks and other advanced

fields of Computer Science. Python had deep focus on code readability & this class will teach you python from basics.

4.3.1 Characteristics :

- It provides rich data types and easier to read syntax than any other programming languages
- It is a platform independent scripted language with full access to operating system API's
- Compared to other programming languages, it allows more run-time flexibility
- It includes the basic text manipulation facilities of Perl and Awk
- A module in Python may have one or more classes and free functions
- Libraries in Python are cross-platform compatible with Linux, Macintosh, and Windows
- For building large applications, Python can be compiled to byte-code
- Python supports functional and structured programming as well as OOP
- It supports interactive mode that allows interacting Testing and debugging of snippets of code
- In Python, since there is no compilation step, editing, debugging and testing is fast.

4.3.2 Applications :

Web Applications :

You can create scalable Web Apps using frameworks and CMS (Content Management System) that are built on Python. Some of the popular platforms for creating Web Apps are: Django, Flask, Pyramid, Plone, Django CMS. Sites like Mozilla, Reddit, Instagram and PBS are written in Python.

Scientific and Numeric Computing :

There are numerous libraries available in Python for scientific and numeric computing. There are libraries like: SciPy and NumPy that are used in general

purpose computing. And, there are specific libraries like: EarthPy for earth science, AstroPy for Astronomy and so on. Also, the language is heavily used in machine learning, data mining and deep learning.

Creating software Prototypes :

Python is slow compared to compiled languages like C++ and Java. It might not be a good choice if resources are limited and efficiency is a must. However, Python is a great language for creating prototypes. For example: You can use Pygame (library for creating games) to create your game's prototype first. If you like the prototype, you can use language like C++ to create the actual game.

Good Language to Teach Programming :

Python is used by many companies to teach programming to kids and newbies. It is a good language with a lot of features and capabilities. Yet, it's one of the easiest language to learn because of its simple easy-to-use syntax.

4.4 SYSTEM DESIGN :

4.4.1 System Architecture :

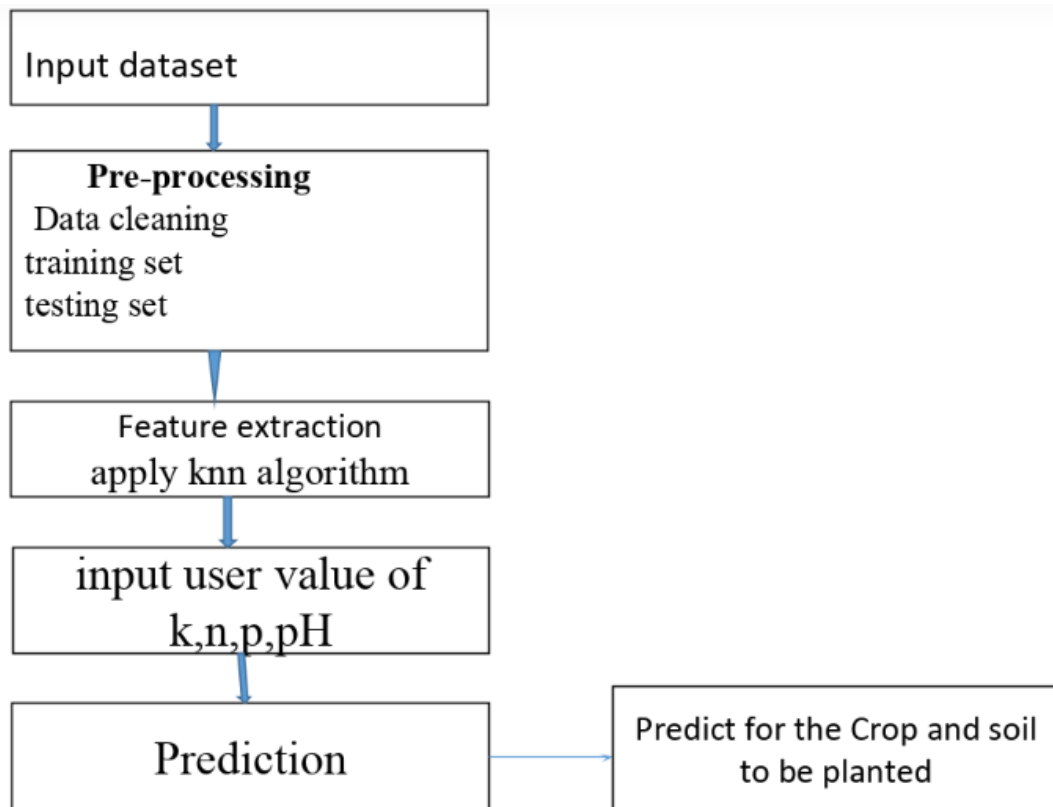


Fig 4.4.1 : SYSTEM ARCHITCTURE

4.4.2 UML Diagrams :

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modeling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems.

The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

4.4.3 Goals :

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modeling Language so that they can develop and exchange meaningful models.
2. Provide extendibility and specialization mechanisms to extend the core concepts.
3. Be independent of particular programming languages and development process.
4. Provide a formal basis for understanding the modeling language.
5. Encourage the growth of OO tools market.
6. Support higher level development concepts such as collaborations, frameworks, patterns and components.
7. Integrate best practices.

4.4.4 Use Case Diagram:

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

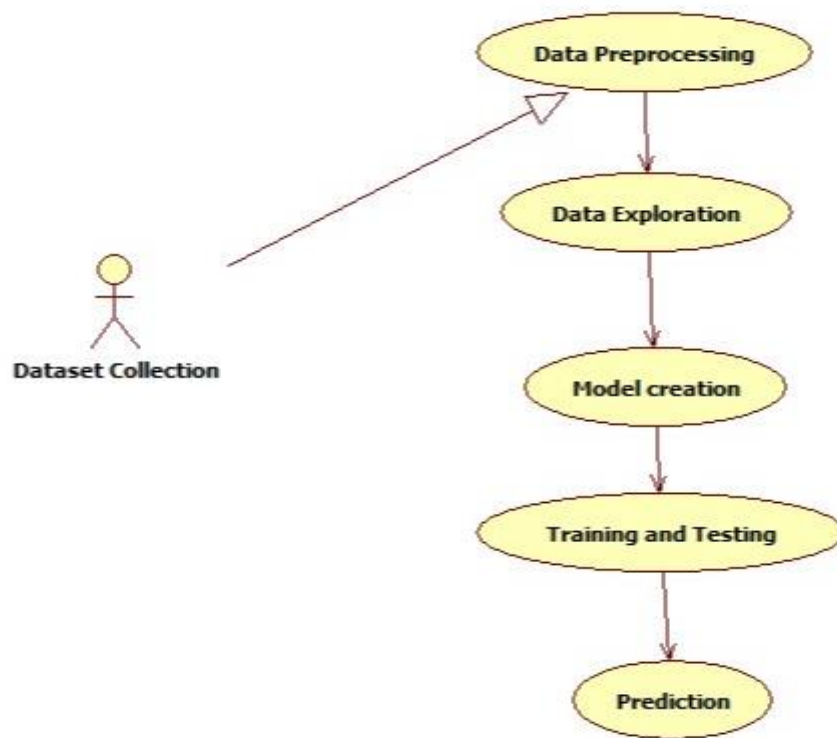


Fig 4.4.4 : USE CASE DIAGRAM

4.4.5 Sequence Diagram :

A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.

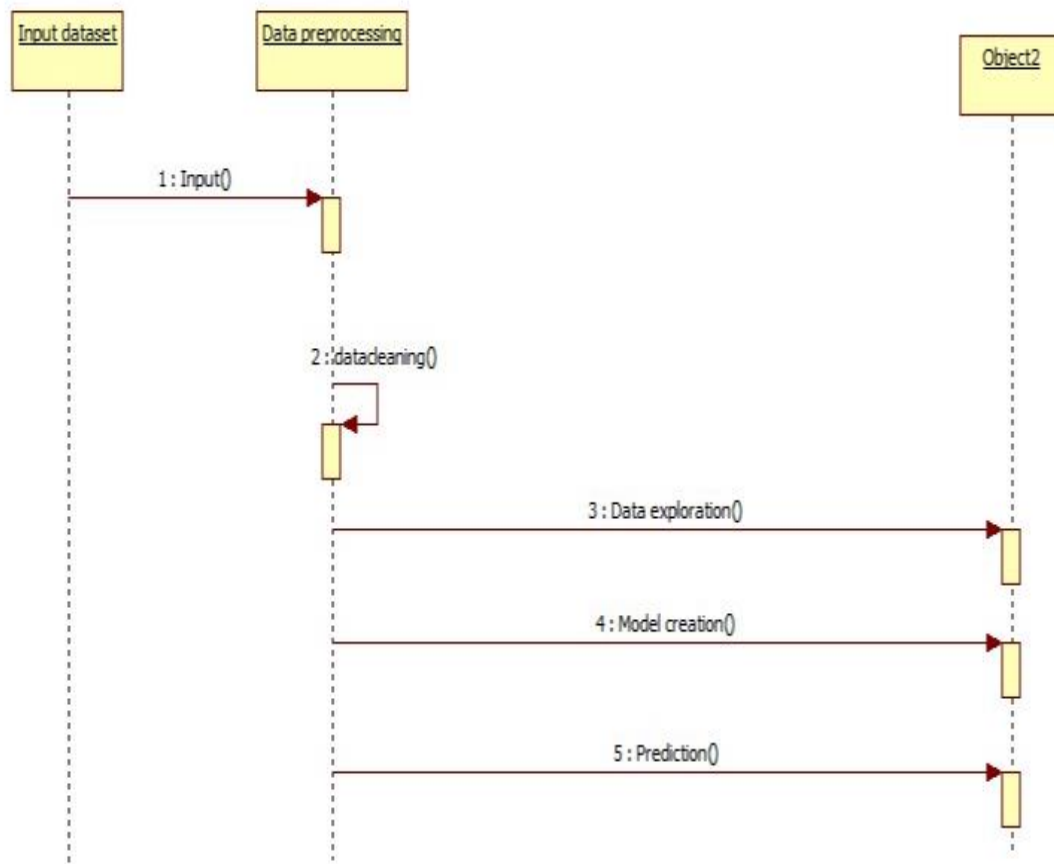


Fig 4.4.5 : SEQUENCE DIAGRAM

4.4.6 Activity Diagram :

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

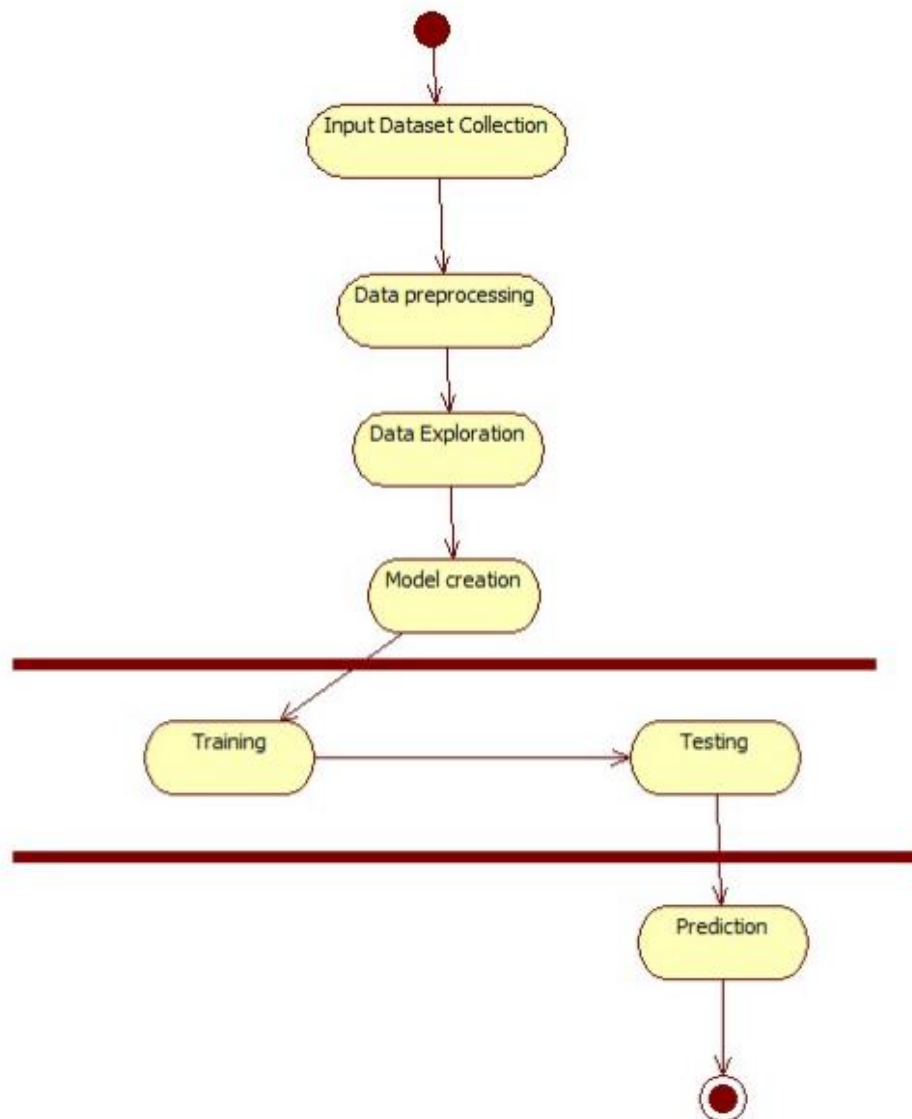


Fig 4.4.6 : ACTIVITY DIAGRAM

CHAPTER – 5

RESULT AND PERFORMANCE ANALYSIS

5.1 MODULE DESCRIPTION :

5.1.1 Data Pre-Processing :

Here the raw data in the crop data is cleaned and the metadata is appended 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 pre-processing, we first load the metadata into this and then this metadata will be attached to the data and replace the converted data with metadata. Then this data will be moved further and remove the unwanted data in the list and it will divide the data into the train and the test data. For this splitting of the data into train and test we need to import `train_test_split` which in the scikit-learn this will help the pre-processed data to split the data into train and test according to the given weight given in the code. The division of the test and train is done in 0.2 and 0.8 that is 20 and 80 percent respectively.

Model Creation:

We create datas into two models:

A) Training model

B) Testing model

The division of the test and train is done in 0.2 and 0.8 that is 20 and 80 percent respectively.

Model evaluation:

We apply the machine learning algorithm for testing part and get the accuracy of this model.

Prediction:

This module based on GUI part. we create a web page using bootstrap. The web page like (enter the value of nitrogen, enter the value of phosphorus, enter the value of potassium). .now we get the data's from user to compare the dataset values .finally it will predict for the Crop and soil to be planted.

Methodology:

First we select one places in google map. Give the value of nitrogen, phosphorus, potassium and pH value. We already trained the dataset. Our value compared to dataset and finally result will displayed what seed we cultivated that particular place.

5.1.2 Data Set Description :

This is the sample data set used in this project. The data in Table I is data used to predict crop yield based on 8 factors. These 8 factors are Garlic, Onion, Orange, Peas, Potato, Rice, Tomato, Sugarcane. We can create a machine learning model and train the model and we can predict the production and from Table II we can predict the amount of fertilizer should be used to get the proper yield the input parameters are the quantity of nitrogen, phosphorus, and the output is the amount of the respective fertilizer should be used. Hear in the input parameters 1, 2, 3, 4, 5, 6,7,8 represents the present in the soil respectively.

Necessary Packages:

1. Numpy
2. Pandas
3. Matplotlib
4. pyplot
5. Scikit-learn
6. Tensorflow
7. Jupyter

5.1.3 Algorithm :

KNN can be used for both classification and regression predictive problems. However, it is more widely used in classification problems in the industry. To evaluate any technique we generally look at 3 important aspects:

1. Ease to interpret output
2. Calculation time
3. Predictive Power

KNN algorithm fails across all parameters of considerations. It is commonly used for its easy of interpretation and low calculation time.

5.2 SOFTWARE TESTING :

General :

In a generalized way, we can say that the system testing is a type of testing in which the main aim is to make sure that system performs efficiently and seamlessly. The process of testing is applied to a program with the main aim to discover an unprecedented error, an error which otherwise could have damaged the future of the software. Test cases which brings up a high possibility of discovering an error is considered successful. This successful test helps to answer the still unknown errors.

Test Case

Testing, as already explained earlier, is the process of discovering all possible weak-points in the finalized software product. Testing helps to counter the working of sub-assemblies, components, assembly and the complete result. The software is taken through different exercises with the main aim of making sure that software meets the business requirement and user-expectations and doesn't fail abruptly. Several types of tests are used today. Each test type addresses a specific testing requirement.

Testing Techniques :

A test plan is a document which describes approach, its scope, its resources and the schedule of aimed testing exercises. It helps to identify almost other test item, the features which are to be tested, its tasks, how will everyone do each task, how much the tester is independent, the environment in which the test is taking place, its technique of design plus the both the end criteria which is used, also rational of choice of theirs, and whatever kind of risk which requires emergency planning. It can be also referred to as the record of the process of test planning. Test plans are usually prepared with signification input from test engineers.

5.2.1 Unit Testing :

In unit testing, the design of the test cases is involved that helps in the validation of the internal program logic. The validation of all the decision branches and internal code takes place. After the individual unit is completed it takes place. Plus it is taken into account after the individual unit is completed before integration. The unit test thus performs the basic level test at its component stage and test the particular business process, system configurations etc. The unit test ensures that the particular unique path of the process gets performed precisely to the documented specifications and contains clearly defined inputs with the results which are expected.

5.2.2 Integration Testing :

These tests are designed to test the integrated software items to determine whether if they really execute as a single program or application. The testing is event driven and thus is concerned with the basic outcome of field. The Integration tests demonstrate that the components were individually satisfaction, as already represented by successful unit testing, the components are apt and fine. This type of testing is specially aimed to expose the issues that come-up by the components combination.

5.2.3 Functional Testing :

The functional tests help in providing the systematic representation that functions tested are available and specified by technical requirement, documentation of the system and the user manual.

5.2.4 System Testing :

System testing, as the name suggests, is the type of testing in which ensure that the software system meet the business requirements and aim. Testing of the configuration is taken place here to ensure predictable result and thus analysis of it. System testing is relied on the description of process and its flow, stressing on pre driven process and the points of integration.

5.2.5 White Box Testing :

The white box testing is the type of testing in which the internal components of the system software is open and can be processed by the tester. It is therefore a complex type of testing process. All the data structure, components etc. are tested by the tester himself to find out a possible bug or error. It is used in situation in which the black box is incapable of finding out a bug. It is a complex type of testing which takes more time to get applied.

5.2.6 Black Box Testing :

The black box testing is the type of testing in which the internal components of the software is hidden and only the input and output of the system is the key for the tester to find out a bug. It is therefore a simple type of testing. A programmer with basic knowledge can also process this type of testing. It is less time consuming as compared to the white box testing. It is very successful for software which are less complex are straight-forward in nature. It is also less costly than white box testing.

5.2.7 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.

5.3 SCREENSHOTS :

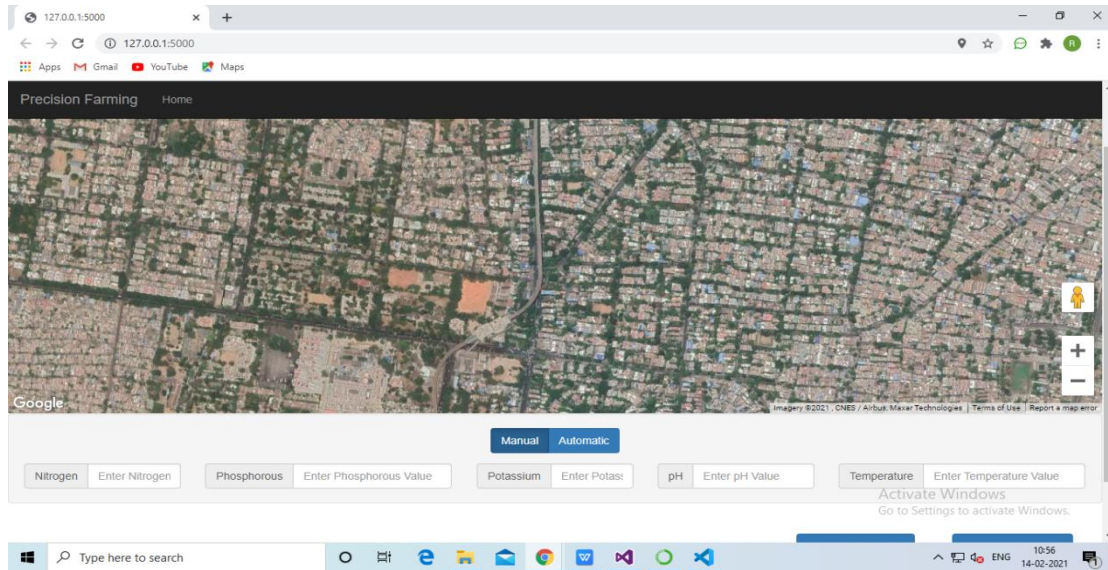


Fig No. 5.3.1 Data Entry Page

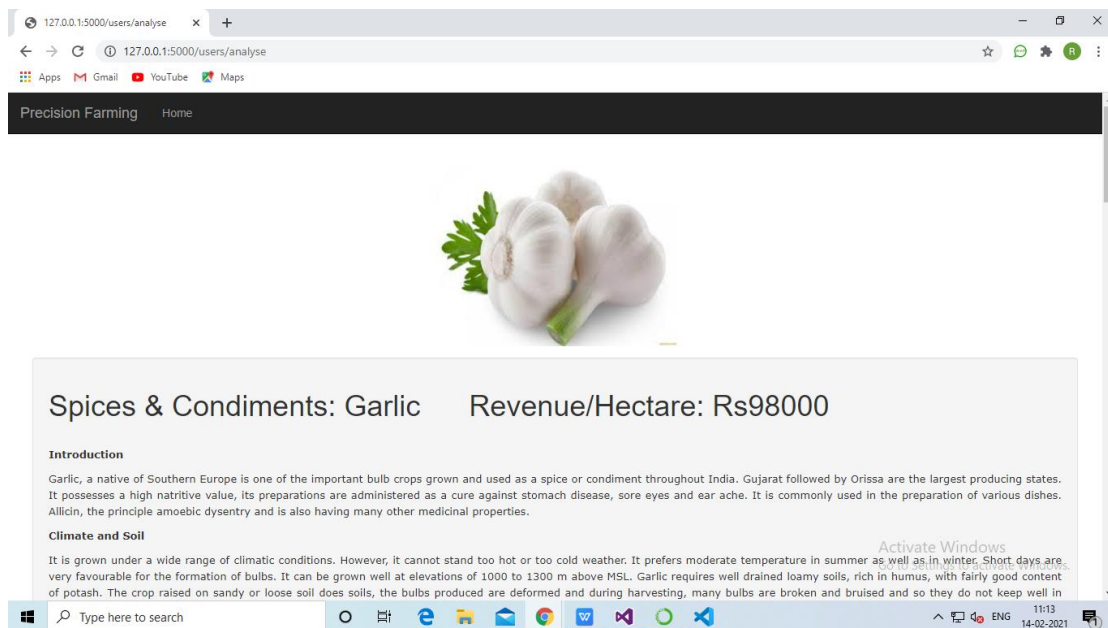


Fig No. 5.3.2 Output for some data values

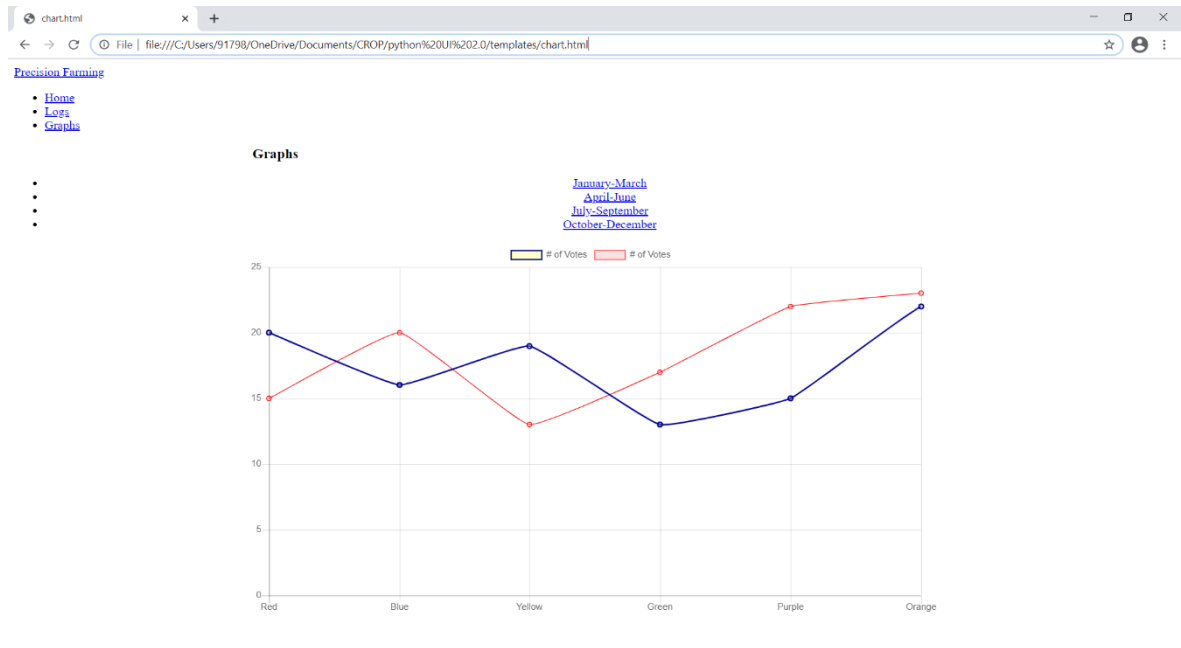


Fig No. 5.3.2 Graph(o/p)

CHAPTER – 6

CONCLUSION AND FUTURE WORK

CONCLUSION :

It is found that by applying different machine learning algorithm and techniques on yields of crops, it is possible to predict crop yield information, which helps to government organization for making better decisions and applying different policies, also it's helpful for farmers to make better plan and increase their production. It's lead to increasing the Countries' overall profit. In our project we found that the accurate prediction of different specified crop yields across different districts will help to farmer. From this farmers will plant different crops in different districts. In the near future, geospatial analysis will be added to our data processing model to improve accuracy and also implement a better geographical data.

FUTURE WORK :

The proposed work presents a proficient yield suggestion framework utilizing classifier models. The framework is versatile as it very well may be utilized to test on various crops. From the yield diagrams the best season of planting, plant development and reaping of plant can likewise be discovered alongside forecast for crops. Proposed system utilizing Machine learning. Choice tree shows execution well, when dataset is having more varieties and gives better outcome to datasets. In this way the framework will help decrease the challenges looked by the ranchers and prevent them from endeavoring suicides. It will go about as a medium to give the ranchers effective data needed to get high return and consequently augment benefits which thus will diminish the self destruction rates and reduce his challenges.

REFERENCES

- [1] Mayank Champaneri, Chaitanya Chandvidkar , Darpan Chachpara, Mansing Rathod, "Crop yield prediction using machine learning" International Journal of Science and Research ,April 2020.
- [2] Pavan Patil, Virendra Panpatil, Prof. Shrikant Kokate, "Crop Prediction System using Machine Learning Algorithms", International Research Journal of Engineering and Technology, Feb 2020.
- [3] Ramesh Medar , Shweta, Vijay S. Rajpurohit, "Crop Yield Prediction using Machine Learning Techniques", 5th International Conference for Convergence in Technology, 2019.
- [4] Trupti Bhange, Swati Shekapure, Komal Pawar, Harshada Choudhari, "Survey Paper on Prediction of Crop yield and Suitable Crop", International Journal of Innovative Research in Science, Engineering and Technology, May 2019.
- [5] E. Manjula, S. Djodiltachoumy, "A Modal for Prediction of Crop Yield "International Journal of Computational Intelligence and Informatics, March 2017.
- [6] Nishit Jain, Amit Kumar, Sahil Garud, Vishal Pradhan, Prajakta Kulkarni, "Crop Selection Method Based on Various Environmental Factors Using Machine Learning", International Research Journal of Engineering and Technology (IRJET), Feb 2017.
- [7] Rakesh Kumar, M.P. Singh, Prabhat Kumar, J.P. Singh, "Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique", 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM),Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, T.N., India., May 2015.
- [8] Rajshekhar Borate., "Applying Data Mining Techniques to Predict Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in India", International Journal of Novel Research in Computer Science and Software Engineering, Vol. 3, Issue 1, pp: (34-37), April 2016.
- [9] D Ramesh, B Vishnu Vardhan, "Analysis of Crop Yield Prediction using Data Mining Techniques", International Journal of Research in Engineering and Technology (IJRET), Vol.4, 2015.
- [10]. S.Veenadhari, Dr Bharat Misra, Dr CD Singh.2019."Machine learning approach for forecasting crop yield based on climatic parameters.".978-1-4799-2352-6/14/\$31.00 ©2014 IEEE.

- [11]. Igor Oliveira, Renato L. F. Cunha, Bruno Silva, Marco A. S. Netto.2018."A Scalable Machine Learning System for PreSeason Agriculture Yield Forecast.".978-1-5386-9156- 4/18/\$31.00 ©2018.
- [12] Neha Rale, Raxitkumar Solanki, Doina Bein, James Andro-Vasko, Wolfgang Bein."Prediction of Crop Cultivation".978-1-7281-0554-3/19/\$31.00©2019 IEEE.
- [13]. Md. Tahmid Shakoor, Karishma Rahman, Sumaiya Nasrin Rayta, Amitabha Chakrabarty.2017."Agricultural Production Output Prediction Using Supervised Machine Learning Techniques".978-1-5386-3831-6/17/\$31.00 ©2017 IEEE.
- [14]. G Srivatsa Sharma, Shah Nawaz Mandal, Shruti Kulkarni, Monica R Mundada, Meeradevi.2018."Predictive Analysis to Improve Crop Yield Using a Neural Network Model".978-1-5386-5314-2/18/\$31.00 ©2018 IEEE.
- [15]. Rashmi Priya, Dharavath Ramesh.2018."Crop Prediction on the Region Belts of India: A Naïve Bayes MapReduce Precision Agricultural Model". 978-1-5386-5314- 2/18/\$31.00 ©2018 IEEE.

APPENDIX

A. PAPER ACCEPTANCE MAIL :

B. PLAGIARISM REPORT :

C. JOURNAL PAPER :

D. SOURCE CODE :

Server.py

```
from flask import Flask , render_template , request , redirect , url_for
from sklearn.neighbors import KNeighborsClassifier
app = Flask(__name__)

_code = ""
@app.route('/')
def index2():
    return render_template('index.html')
@app.route('/Register')
def Register():
    return render_template('register.html')

@app.route("/users/index", methods=["GET","POST"])
def main():
    return render_template('index.html')

@app.route("/users/graph", methods=["GET","POST"])
def login():
    return render_template('chart.html')

@app.route("/users/logs", methods=["GET","POST"])
def logs():
    return render_template('logs.html')

@app.route("/users/analyse", methods=["POST"])
def analyse():
    if(request.method == "POST"):
        import pandas as pd
        import numpy as np
        # import os
```

```

    optimum = pd.read_excel(r"D:\Shahana\ML Project Correct\CROP\python UI
2.0\optimum2.xlsx", 'newData')

    price = pd.read_excel(r"D:\Shahana\ML Project Correct\CROP\python UI 2.0\
optimum2.xlsx", 'pricePerhr')

    optimum['N'] = optimum.N.astype(float)
    optimum['P'] = optimum.P.astype(float)
    optimum['K'] = optimum.K.astype(float)
    optimum['TEMPERATURE'] = optimum.TEMPERATURE.astype(float)
    X = optimum.drop("CLASS",axis=1)
    y = optimum.CLASS

from sklearn.neighbors import KNeighborsClassifier
clf = KNeighborsClassifier(n_neighbors=3)
clf.fit(X,y)

print(request.form.get('Potassium'))

if(request.form.get('Potassium') == None):

    pred = pd.read_excel(r'D:\Shahana\ML Project Correct\CROP\python UI 2.
0\optimum2.xlsx', 'Sheet3')

    prediction = clf.predict(pred)
    print(prediction)

    optimum = optimum[optimum['CLASS'] != prediction[0]]
    X = optimum.drop("CLASS",axis=1)
    y = optimum.CLASS
    clf = KNeighborsClassifier(n_neighbors=3)
    clf.fit(X,y)
    prediction1 = clf.predict(pred)
    print(prediction1)

    optimum = optimum[optimum['CLASS'] != prediction1[0]]

```

```

X = optimum.drop("CLASS",axis=1)
y = optimum.CLASS
clf = KNeighborsClassifier(n_neighbors=3)
clf.fit(X,y)
prediction2 = clf.predict(pred)
print(prediction2)
p1 = prediction1[0]
p2 = prediction2[0]
p1 = p1 -1
p2 = p2 -1
# print()

if(prediction == 7):
    return render_template('crops.html' , crop="TOMATO" , crop1=prediction
1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[6] ,price1=price["Price/hr"].il
oc[p1] , price2=price["Price/hr"].iloc[p2])
elif(prediction == 1):
    return render_template('crops.html' , crop="GARLIC" , crop1=prediction1
[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[[0]] ,price1=price["Price/hr"].il
oc[p1] , price2=price["Price/hr"].iloc[p2])
elif(prediction == 2):
    return render_template('crops.html' , crop="ONION" , crop1=prediction1[
0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[[1]] , price1=price["Price/hr"].il
oc[p1] , price2=price["Price/hr"].iloc[p2])
elif(prediction == 3):
    return render_template('crops.html' , crop="ORANGE" , crop1=prediction
1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[[2]] , price1=price["Price/hr"]
.iloc[p1] , price2=price["Price/hr"].iloc[p2])
elif(prediction == 4):
    return render_template('crops.html' , crop="PEAS" , crop1=prediction1[0]
, crop2=prediction2[0] , price=price["Price/hr"].iloc[[3]] , price1=price["Price/hr"].ilo
c[p1] , price2=price["Price/hr"].iloc[p2])
elif(prediction == 5):

```

```

        return render_template('crops.html' , crop="POTATO" , crop1=prediction
1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[[4]] , price1=price["Price/hr"]
.iloc[p1] , price2=price["Price/hr"].iloc[p2])
    elif(prediction == 6):
        return render_template('crops.html' , crop="RICE" , crop1=prediction1[0]
, crop2=prediction2[0] , price=price["Price/hr"].iloc[[5]] , price1=price["Price/hr"].iloc
[p1] , price2=price["Price/hr"].iloc[p2])
    elif(prediction == 8):
        return render_template('crops.html' , crop="SUGARCANE" , crop1=predi
ction1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[[7]] , price1=price["Pric
e/hr"].iloc[p1] , price2=price["Price/hr"].iloc[p2])
    else:
        return "no"

else:
    potassium = request.form.get('Potassium')
    phosphorous = request.form.get('Phosphorous')
    nitrogen = request.form.get('Nitrogen')
    pH = request.form.get('pH')
    temperature = request.form.get('Temperature')

    columns = ['N','P','K','pH','TEMPERATURE']
    values = np.array([ nitrogen ,phosphorous ,potassium ,  pH , temperature])
    pred = pd.DataFrame(values.reshape(-1, len(values)),columns=columns)
    # print(pred.dtype)
    print(pred)

    prediction = clf.predict(pred)
    print(prediction)

    optimum = optimum[optimum['CLASS'] != prediction[0]]
    X = optimum.drop("CLASS",axis=1)
    y = optimum.CLASS
    clf = KNeighborsClassifier(n_neighbors=3)

```

```

clf.fit(X,y)
prediction1 = clf.predict(pred)
print(prediction1)

optimum = optimum[optimum['CLASS'] != prediction1[0]]
X = optimum.drop("CLASS",axis=1)
y = optimum.CLASS
clf = KNeighborsClassifier(n_neighbors=3)
clf.fit(X,y)
prediction2 = clf.predict(pred)
print(prediction2)

p1 = prediction1[0]
p2 = prediction2[0]
p1 = p1 -1
p2 = p2 -1
# print()

if(prediction == 7):
    return render_template('crops.html' , crop="TOMATO" , crop1=prediction1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[6] ,price1=price["Price/hr"].iloc[p1] , price2=price["Price/hr"].iloc[p2])
elif(prediction == 1):
    return render_template('crops.html' , crop="GARLIC" , crop1=prediction1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[0] ,price1=price["Price/hr"].iloc[p1] , price2=price["Price/hr"].iloc[p2])
elif(prediction == 2):
    return render_template('crops.html' , crop="ONION" , crop1=prediction1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[1] , price1=price["Price/hr"].iloc[p1] , price2=price["Price/hr"].iloc[p2])
elif(prediction == 3):
    return render_template('crops.html' , crop="ORANGE" , crop1=prediction1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[2] , price1=price["Price/hr"].iloc[p1] , price2=price["Price/hr"].iloc[p2])

```

```

        elif(prediction == 4):
            return render_template('crops.html' , crop="PEAS" , crop1=prediction1[0]
, crop2=prediction2[0] , price=price["Price/hr"].iloc[3] , price1=price["Price/hr"].iloc[
p1] , price2=price["Price/hr"].iloc[p2])
        elif(prediction == 5):
            return render_template('crops.html' , crop="POTATO" , crop1=prediction
1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[4] , price1=price["Price/hr"].i
loc[p1] , price2=price["Price/hr"].iloc[p2])
        elif(prediction == 6):
            return render_template('crops.html' , crop="RICE" , crop1=prediction1[0]
, crop2=prediction2[0] , price=price["Price/hr"].iloc[5] , price1=price["Price/hr"].iloc[
p1] , price2=price["Price/hr"].iloc[p2])
        elif(prediction == 8):
            return render_template('crops.html' , crop="SUGARCANE" , crop1=predi
ction1[0] , crop2=prediction2[0] , price=price["Price/hr"].iloc[7] , price1=price["Price/
hr"].iloc[p1] , price2=price["Price/hr"].iloc[p2])
        else:
            return "no"
    # render_template('index.html')
else:
    # return render_template('index.html')
    pass

```

```

if (__name__ == "__main__"):
    app.run()
Index.html
<!DOCTYPE html>
<script type="text/javascript">
var centerlat , centerlong;
var staticMapUrl = "https://maps.googleapis.com/maps/api/staticmap";
function myMap() {
var coordinateslt = [];
var coordinateslg = [];
var loclat , loclng ;

```

```

var coordinates = [];
var num = coordinates.length;
var map,location,mapCanvas,mapOptions;

if (navigator.geolocation) {
    navigator.geolocation.getCurrentPosition(function(position) {
        loclat = position.coords.latitude;
        loclng = position.coords.longitude;
        location = new google.maps.LatLng(loclat , loclng);
        mapCanvas = document.getElementById("map");
        mapOptions = {center: location, zoom: 16};
        map = new google.maps.Map(mapCanvas,mapOptions);
        map.setMapTypeId(google.maps.MapTypeId.SATELLITE);

        function setpoly(){

for(var i=0;i<num;i++)
    {
        coordinates.push(new google.maps.LatLng(coordinateslt[i] , coordinateslg[i]));
    }
    coordinates.pop();
    map.setCenter(coordinates[num-1]);

    flightPath = new google.maps.Polygon({
        path:coordinates,
        strokeColor: "#0000FF",
        strokeOpacity: 0.8,
        strokeWeight: 2,
        fillColor: "#0000FF",
        fillOpacity: 0.4
    });
    flightPath.setMap(map);
    google.maps.event.clearListeners(map, "rightclick");

```

```

        google.maps.event.clearListeners(map, "click");
    }

    google.maps.event.addListener(map, "rightclick", function(event) {
        var lat = event.latLng.lat();
        var lng = event.latLng.lng();
        coordinateslt.push(lat);
        coordinateslg.push(lng);
        num = num + 1;
        setpoly();
        centerlong = lng;
        centerlat = lat;
        staticMapUrl += "?center=" + centerlat + "," + centerlong;
        staticMapUrl += "&size=220x300";
        staticMapUrl += "&zoom=" + 18;
        staticMapUrl += "&maptype=" + google.maps.MapTypeId.SATELLITE;

    });

    google.maps.event.addListener(map, "click", function(event) {
        var lat = event.latLng.lat();
        var lng = event.latLng.lng();
        coordinateslt.push(lat);
        coordinateslg.push(lng);
        num = num + 1;
    });

    });
}
}

```