**CHAPTER 1**

**INTRODUCTION**

**1.1PROBLEM DEFINITION:**

Global temperatures rose above pre-industrial levels by + 0.85 °C in the last century, and are predicted to exceed + 2 °C this century. There are aspirations to limit this temperature rise by reducing anthropogenic [greenhouse gas](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/greenhouse-gases) emissions, but current [global warming](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/global-warming) trends are expected to lead to a greater intensity, frequency and severity of [droughts](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/drought). Higher temperature and increased rainfall variability will reduce yields of major crops such as maize, wheat, paddy, blackgram, greengram and redgram in spite of the benefits for plants from increased atmospheric CO2.

Rainfed areas supply ca. 58% of global food production and play an important role in food security. Rice is one of the major crops grown and consumed in rainfed areas, and rainfed cultivation accounts for about 25% of global rice production. Due to its dependence on climate, rainfed rice cultivation is vulnerable to changes in temperature and rainfall. Warm temperature (optimal range 20 °C–30 °C) and high rainfall (optimal range 1500 mm–2000 mm) generally increase growth rates of different plants, and hence yield. By contrast, very high temperatures (> 35 °C) induce heat stress and affect plant physiological processes, leading to [spikelet](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/spadix) sterility, non-viable pollen and reduced [grain quality](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/grain-quality). Drought, on the other hand, reduces plant [transpiration](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/transpiration) rates and may result in [leaf rolling](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/leaf-rolling) and drying, reduction in leaf expansion rates and plant [biomass](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/biomass), immobilisation of [solutes](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/solutes) and increased heat stress of leaves.

Climate is the primary factor driving locations for rainfed rice cultivation and rice yields. Hence changes in climate, such as those projected to occur in the future, particularly those related to increased variability in rainfall, could result in some areas becoming climatically unsuitable for cultivating rainfed rice, or at least reduce crop yields. Statistical models have been used to map crop production in relation to climate, and to project changes in the suitability of cultivation for a wide variety of crops including cereals, and fruit. Climate envelope models (CEMs) have been used at regional scales to map distributions of crops in relation to climate variables and, by incorporating outputs from future climate change scenarios, to make projections about changes in the suitability of cropping areas. Generally, outputs of CEMs are expressed in terms of spatial (usually gridded) maps of probabilities of occurrence of the crop under study, with declines in probability under future climate change implying decreasing suitability for growing crops. CEM outputs can be used to identify regions that may become climatically unsuitable in the future, and highlight vulnerable areas where crops are most at risk from the detrimental impacts of climate change. This mapping approach can be used at regional scales to guide policy makers in their choice of adaptation strategies, such as breeding new [cultivars](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/cultivar) that can cope with the predicted climate change, developing irrigation infrastructure or shifting to new [cropping systems](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/cropping-system).

In this study, we examine changes in climate suitability of rainfed rice cultivation in India, to highlight areas at risk from future climate changes. It is important to study rainfed rice cultivation here because India is the world's second largest producer of rice, of which a substantial amount is grown under rainfed conditions during the *Kharif* (i.e. summer monsoon season). Any detrimental impacts of climate would have major consequences for food security from local to global levels. Moreover, the majority of Indian farmers cultivating rainfed rice are smallholders, whose local [livelihoods](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/livelihood) are highly vulnerable to climate changes and since 1980, the number of smallholder farmers in India increased by ~ 87% to almost 66 million in 2017–18. In addition, the agricultural sector in India employs almost half of the labour force of the country, so any changes in rice cultivation are likely to have considerable social impacts.

We use multiple CEMs and BRTs and extent of rainfed rice cultivation in relation to four climate variables during the main summer monsoon growing season Modelling continuous data, i.e. extent of rainfed rice using boosted regression trees, as well as categorical occurrence data using CEMs, allowed us to map changes in the suitability of rainfed rice growing areas, as well as to quantify changes in the absolute area available for rainfed rice cultivation. Our study has three main aims. First, we examine whether the occurrence and extent of current-day rainfed rice cultivation can be modelled successfully using climatic variables derived from temperature and precipitation during the summer monsoon, and whether CEM and BRT model outputs agree in terms of which areas are climatically most suitable for growing rainfed rice. Second, we assess whether the models agree on which climate variables are important predictors of rainfed rice cultivation; we hypothesise here that rainfall-derived variables will be more important than temperature in this respect. Finally, we map future changes in the climate suitability of areas where rainfed rice is currently cultivated, and identify risk areas that our models project to possibly become climatically unsuitable for rainfed rice cultivation by 2050.

* 1. **EXISTING SYSTEM:**

The Existing Algorithms provides most important benefits are interpretability. Generally, agricultural department stores the information about the different factors that leads to crop yield. It mainly focuses on   production increment and productivity of agricultural crops using environment friendly science and technology while ensuring increased net farm income to the farmers through various schemes, programmes and welfare measures. However, does not investigate the root cause for the decrease in crop yield/agriculture risk. They store the information manually. This information is not used for future analysis. So, further prediction from previous data is required.

**1.2.1DISADVANTAGES OF EXISTING SYSTEM:**

The existing system results in the following drawbacks:

* It does not predict the future scope of the yield.

* 1. **PROPOSED SYSTEM:**

Our project processes the data collected from records of agriculture departments which are stored manually through data visualization tools like weka/tableau/R tool. The result will be presented in a graphical format for easy understanding of user. By using this result the agricultural department will make a decision for future crop production. This will minimizing risk increasing the crop production.

We are using tool kit in the Data Analyzer of Microsoft Excel. We are deploying the Residual plots by using regression function which is provided in the tool kit.

* + 1. **ADVANTAGES OF PROPOSED SYSTEM:**

The proposed system results in the following advantages:

* Proposed system gives the esstimted value of yield depending upon values of rainfall and area of crop
* It is easy to take a decision on future crop production.

**1.4 SOFTWARE REQUIREMENTS SPECIFICATION:**

**Software Requirements:**

Operating system : Windows95/98/2000/XP.

Data Mining Tool : R-Tool.

Database : Agriculture Data Sets(including yield,productivity,area ),

And Rainfall Data

**Hardware Requirements:**

Processor : Pentium-III.

Speed : 1.1 GHz.

RAM : 2 GB (or) Higher.

Hard Disk : Minimum of 1GB.

**1.5 SYSTEM REQUIREMENT SPECIFICATION**

**FUNCTIONAL AND NON-FUNCTIONAL REQUIREMENTS**

A **System Requirements Specification (SRS) –** a requirements specification for a software system – is a complete description of the behavior of a system to be developed. It includes a set of use cases that describe all the interactions the users will have with the software. Use cases are also known as functional requirements. In addition to use cases, the SRS also contains non-functional (or supplementary) requirements. Non – functional requirements are requirements which impose constraints on the design or implementation (such as performance engineering requirements, quality standards, or design constraints).

**1.5.1FUNCTIONAL REQUIREMENTS**

**LEAST SQUARES METHOD:**

The Regression analysis tool performs linear regression analysis by using the "least squares" method to fit a line through a set of observations. You can analyze how a single dependent variable is affected by the values of one or more independent variables. For example, you can analyze how an athlete's performance is affected by such factors as age, height, and weight. You can apportion shares in the performance measure to each of these three factors, based on a set of performance data, and then use the results to predict the performance of a new, untested athlete.

The Regression tool uses the worksheet function LINEST.

**Regression dialog box**

**Input Y Range**   Enter the reference for the range of dependent data. The range must consist of a single column of data.

**Input X Range**   Enter the reference for the range of independent data. Microsoft Office Excel orders independent variables from this range in ascending order from left to right. The maximum number of independent variables is 16.

**Output Range**   Enter the reference for the upper-left cell of the output table. Allow at least seven columns for the summary output table, which includes an Anova table, coefficients, standard error of y estimate, r2 values, number of observations, and standard error of coefficients.

**LINEST:**

Calculates the statistics for a line by using the “least squares" method to calculate a straight line that best fits your data, and then returns an array that describes the line. You can also combine LINEST with other functions to calculate the statistics for other types of models that are linear in the unknown parameters, including polynomial, logarithmic, exponential, and power series. Because this function returns an array of values, it must be entered as an array formula.

The accuracy of the line calculated by LINEST depends on the degree of scatter in your data. The more linear the data, the more accurate the LINEST model. LINEST uses the method of least squares for determining the best fit for the data. When you have only one independent x-variable, the calculations for m and b are based on the following formulas:

EquationEquation

Where x and y are sample means, i.e., x = AVERAGE (known\_x's) and y = AVERAGE (known\_y's).

**1.5.2 Non-Functional Requirements:**

In systems engineering and requirements engineering, a **non-functional requirement** is a requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors. This should be contrasted with functional requirements that define specific behavior or functions. The plans for implementing non-functional requirements are detailed in the system architecture. The non-functional requirements are **“system shall be <requirement>”**. Non-functional requirements are often called **qualities** of a system.

The following are the Non functional requirements for our system:

**Availability:**

A system's availability, or "uptime," is the amount of time that it is operational and available for use. User can be access system by using executable file or short cut.

**Accessibility:**

Specifies how well the software utilizes scarce resources: CPU cycles, disk space, memory, bandwidth, etc. Since this is developed as pc client application this will able to access only single user at a time.

**Portability:**

Portability specifies the ease with which the software can be installed on all necessary platforms, and the platforms on which it is expected to run. By using appropriate server versions released for different platforms our project can be easily operated on any operating system, hence can be said highly portable.

**Efficiency:**

Specifies about the data classification and predict the suitable results pertaining to the symptoms based on the nutrition values of N-P-K. The software utilization can be effectively used in terms of memory, disk space.

**Scalability:**

Software that is scalable has the ability to handle a wide variety of system configuration sizes. The nonfunctional requirements should specify the ways in which the system may be expected to scale up (by increasing hardware capacity, adding machines etc.). Our system can be easily expandable. Any additional requirements such as hardware or software which increase the performance of the system can be easily added.

**Usability:**

For every type of soil nutrients, can apply this algorithm and get the desired results.

**Performance:**

The performance constraints specify the datasets of the nutrient values. We classify the nutrient values based on plant symptoms. After the classification mainly observe the rice diseases in growth stages.

* Based on the different levels the diseases are identified.
* The normal values and critical values are taken for prediction by using regression function.

**CHAPTER 2**

**SYSTEM ANALYSIS**

**2.1 FLOW ORIENTED MODELING:**

**2.1.1 Data Flow Diagram (DFD):**

A Data Flow Diagram (DFD) is also known as a Process Model.  Process Modeling is an analysis technique used to capture the flow of inputs through a system (or group of processes) to their resulting output.  The model is fairly simple in that there are only four types of symbols – process, dataflow, external entity, data store.

Process Modeling is used to visually represent what a system is doing.  It is much easier to look at a picture and understand the essence than to read through verbiage describing the activities.  System Analyst after talking with various users will create DFD diagrams and then show them to users to verify that their understanding is correct.  The process models can be created to represent an existing system as well as a proposed system.

 The following are some of the symbols in Process Modeling represents.

* Process
* Data Object
* Data Store
* External entity

|  |
| --- |
| **Process:**  An activity or a function that is performed for some specific reason; can be manual or Computerized; ultimately each process should perform only one activity.  **Symbol:** |
| **Data Flow:**  Single piece of data or logical collection of information like a bill.  **Symbol:** |  |
| **Data Store:**  Collection of data that is permanently stored.  **Symbol:**      http://members.tripod.com/%7Emyyee/cs457/datastore.gif  **External Entity:**  A person, organization, or system that is external to the system but intersects with it.  **Symbol:** |  |

The following are the levels of Data Flow Diagrams:

**Level 0 DFD:**

* The level 0 DFD (also known as the context level DFD) is the simplest DFD.
* The outermost level (Level 0) is concerned with how the system interacts with the outside world.
* This level basically represents the input and output of the entire system.

**Level 1 DFD:**

* The basic modules of the system are represented in this phase and how data moves through different module is shown.
* The level 1 DFD provides a high level view of the system that identifies the major processes and data stores.

**Level 2 DFD and other level of DFD:**

* Each process from level 1 is exploded even more into processes. This decomposition continues for each level.
* The number of levels possible depends on the complexity of the system.

**DFD Principles**

* The general principle in Data Flow Diagramming is that a system can be decomposed into subsystems, and subsystems can be decomposed into lower level subsystems, and so on.
* Each subsystem represents a process or activity in which data is processed. At the lowest level, processes can no longer be decomposed.
* Each 'process' (and from now on, by 'process' we mean subsystem and activity) in a DFD has the characteristics of a system.
* Just as a system must have input and output (if it is not dead), so a process must have input and output.
* Data enters the system from the environment; data flows between processes within the system; and data is produced as output from the system.

**Salient Features of DFD’S**

1. The DFD shows flow of data, not of control loops and decision are controlled considerations do not appear on a DFD.
2. The DFD does not indicate the time factor involved in any process whether the dataflow take place daily, weekly, monthly or yearly.
3. The sequence of events is not brought out on the DFD.

**Constructing a DFD:**

Several rules of thumb are used in drawing DFD’S:

1. Process should be named and numbered for an easy reference. Each name should be representative of the process.

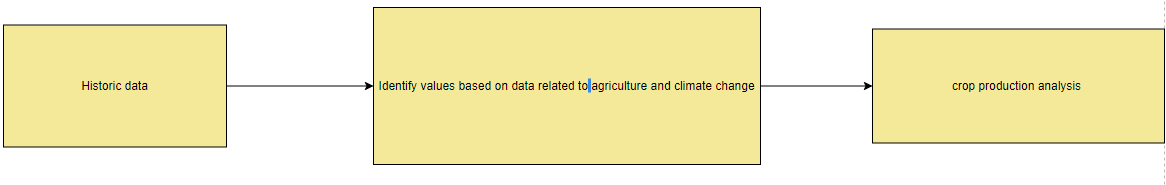
2. The direction of flow is from top to bottom and from left to right. Data traditionally flow from source to the destination although they may flow back to the source. One way to indicate this is to draw long flow line back to a source. An alternative way is to repeat the source symbol as a destination. Since it is used more than once in the DFD it is marked with a short diagonal.

3. The names of data stores and destinations are written in capital letters. Process and dataflow names have the first letter of each work capitalized.

4. A DFD typically shows the minimum contents of data store. Each data store should contain all the data elements that flow in and out.

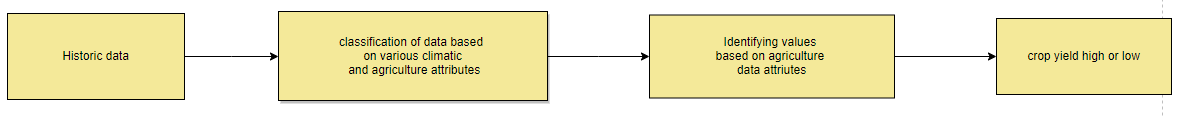
**Data Flow Diagrams:**

**Level-0 DFD:**

****

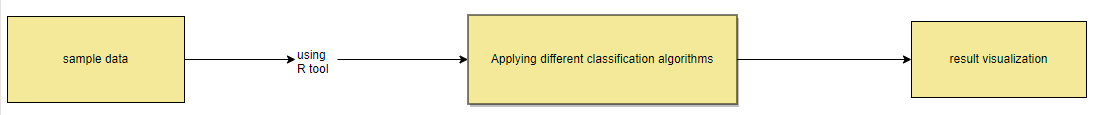
**Fig 2.1: Level-0 Diagram for Identifying Data.**

**Level-1 DFD:**

****

**Fig 2.2: Level-1 Diagram for Classification of Data.**

**Level-2 DFD:**

****

**Fig 2.3: Level 2 Diagram for Applying Regression Algorithm using R tool.**

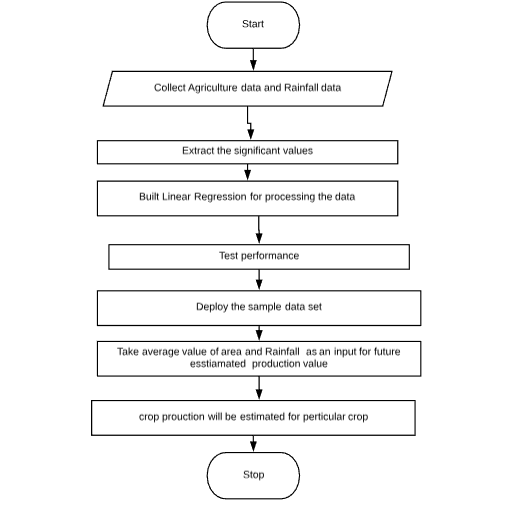
**2.1.2 Flow Charts:**

A **flowchart** is a type of diagram that represents an algorithm or process, showing the steps as boxes of various kinds, and their order by connecting them with arrows. This diagrammatic representation illustrates a solution to a given problem. Process operations are represented in these boxes, and arrows; rather, they are implied by the sequencing of operations. Flowcharts are used in analyzing, designing, documenting or managing a process or program in various fields.

A flow chart diagram typically consists of shapes such as ovals, diamonds and boxes that contain words relating to a certain issue and lines with arrows connecting boxes to one another. Flowcharts are used in designing and documenting complex processes or programs. Like other types of diagrams, they help visualize what is going on and thereby help the viewer to understand a process, and perhaps also find flaws, bottlenecks, and other less-obvious features within it. There are many different types of flowcharts, and each type has its own repertoire of boxes and notational conventions. The two most common types of boxes in a flowchart are:

* a processing step, usually called *activity*, and denoted as a rectangular box
* a decision, usually denoted as a diamond.

**Flow chart diagram:**



**Fig 2.4: Flow Chart Diagram for identifying rice diseases according to the growth stages based on nutritional disorders.**

The Sampling analysis tool creates a sample from a population by treating the input range as a population. When the population is too large to process or chart, you can use a representative sample. You can also create a sample that contains only the values from a particular part of a cycle if you believe that the input data is periodic. For example, if the input range contains quarterly sales figures, sampling with a periodic rate of four places the values from the same quarter in the output range.

**Sampling dialog box**

**Input Range**   Enter the references for the range of data that contains the population of values that you want to sample. Microsoft Office Excel draws samples from the first column, then the second column, and so on.

**Labels**   Select if the first row or column of your input range contains labels. Clear if your input range has no labels. Excel generates the appropriate data labels for the output table.

**Sampling Method**   Click **Periodic** or **Random** to indicate the sampling interval that you want.

**Period**   Enter the periodic interval at which you want the sampling to take place. The *period*-th value in the input range and every *period*-th value thereafter are copied to the output column. Sampling stops when the end of the input range is reached.

**Number of Samples**   Enter the number of random values that you want in the output column. Each value is drawn from a random position in the input range, and any number can be selected more than once.

**Output Range**   Enter the reference for the upper-left cell of the output table. Data is written in a single column below the cell. If you select **Periodic**, the number of values in the output table is equal to the number of values in the input range, divided by the sampling rate. If you select **Random**, the number of values in the output table is equal to the number of samples.

**New Worksheet Ply**   Click to insert a new worksheet in the current workbook and paste the results starting at cell A1 of the new worksheet. To name the new worksheet, type a name in the box.

**New Workbook**   Click to create a new workbook in which results are added to a new worksheet.

|  |  |
| --- | --- |
| **CHAPTER 3**  **SYSTEM IMPLEMENTATION**  **3.1. R- TOOL:**  **3.1.1. Introduction**  **R** is a programming language and free software environment for statistical computing and graphics supported by the **R** Foundation for Statistical Computing. The **R** language is widely used among statisticians and data miners for developing statistical software and data analysis. A number of additional modules and options are also available.  **Main 3.1.2. Main features** of R-tool include :   * R provides a large, coherent and integrated collection of tools for data analysis. * R has an effective data handling and storage facility. * suite of operators for calculations on arrays, lists, vectors and matrices. * graphical facilities for data analysis.   **3.1.3. R-tool download and installation:** To Install R:  * 1. Open an internet browser and go to [www.r-project.org](http://www.r-project.org/).   2. Click the "download R" link in the middle of the page under "Getting Started."   3. Select a CRAN location (a mirror site) and click the corresponding link.   4. Click on the "Download R for Windows" link at the top of the page.   5. Click on the "install R for the first time" link at the top of the page.   6. Click "Download R for Windows" and save the executable file somewhere on your computer.  Run the .exe file and follow the installation instructions.   7. Now that R is installed, you need to download and install RStudio.  To Install RStudio  * 1. Go to [www.rstudio.com](http://www.rstudio.com/) and click on the "Download RStudio" button.   2. Click on "Download RStudio Desktop."   3. Click on the version recommended for your system, or the latest Windows version, and save the executable file.  Run the .exe file and follow the installation instructions.  To Install the SDSFoundations Package  * 1. Download [SDSFoundations](https://preview.edx.org/c4x/UTAustinX/UT.7.01x/asset/SDSFoundations_1.1.zip) to your desktop (make sure it has the ".zip" extension).   2. Open RStudio.   3. Click on the Packages tab in the bottom right window.   4. Click "Install."   5. Select install from "Package Archive File."   6. Select the SDSFoundations package file from your desktop.   7. Click install. You are done! You can now delete the SDSpackage file from your desktop.   8. **Start the R-tool**   **CHAPTER 4**  **SOURCE CODE**  data<- read.csv("H:\\renuka\\major project\\finaldata\\paddy.csv")  data  str(data)  install.packages("GGally")  library(GGally)  fit2 <- lm(ProdinMTs ~Areainha + Rainfall , data = data)  summary(fit2)  M<-c("A Konduru","Agiripalli","Avanigadda","Bantumilli","Bapulapadu","Challapalli","Chandarlapadu","Chatrai","G.Konduru","Gampalagudem","Gannavaram", "Ghatasala","Gudiwada","Gudlavalleru","Guduru","Ibrahimpatnam","Jaggaiahpet","Kaikuluru"," Kalidindi"," Kanchikacherla","Kankipadu","Koduru" ,"Kruthivennu","Machilipatnam", "Mandavalli", "Mopidevi", "Movva", " Mudinepalli","Musunuru","Mylavaram","Nagayalanka"," Nandigama","Nandivada","Nuzvid"," Pamaruu","Pamidimukkala","Pedana"," Pedaparapudi","Penamaluru","Penuganchiprolu","Reddigudem", "Thotlavalluru","Tiruvuru","Unguturu","Vatsavai","Veerullapadu","Vijayawada Rural","Vijayawada Urban","Vissannapet","Vuyyuru")  table<-c(data$Areainha + data$Rainfall)  barplot(table,names.arg=M,xlab =" Areainha + Rainfall",ylab = "ProdinMTs",col="blue",main="Plot",border="blue", cex.names = 0.65,  cex.lab = 0.65,las=2)  predict(fit2,data.frame(Areainha=3000,Rainfall=500.65))  **CHAPTER 5**  **SCREENS**  **C:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (450).png**  **Fig 5.1: Data analysis Graph for cotton and estimated yield value in future.**  **C:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (452).png**  **Fig 5.2: Data analysis Graph for paddy(kharif) and estimated yield in future**  **C:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (452).png**  **Fig 5.3: Data analysis Graph for paddy(raby) and estimated yield in future**  **C:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (453).png**  **Fig 5.4: Data analysis Graph for greengram(raby) and estimated yield in futureC:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (455).png**  **Fig 5.5: Data analysis Graph for Maize(raby) and estimated yield in future**  **C:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (456).png**  **Fig 5.6: Data analysis Graph for redgram(kharif) and estimated yield in future**  **C:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (457).png**  **Fig 5.7: Data analysis Graph for blacgram(raby) and estimated yield in future**  **C:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (458).png**  **Fig 5.8: Data analysis Graph for blacgram(kharif) and estimated yield in future**  **C:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (459).png**  **Fig 5.9: Data analysis Graph for greengram(kharif) and estimated yield in future**  **C:\Users\welcome\Pictures\Saved Pictures\Screenshots\Screenshot (470).png**  **Fig 5.10: Data analysis Graph for gruondnut(kharif) and estimated yield in future**  **CHAPTER 6**  **WATER MANAGEMENT TECHNIQUES**  Since water is essential to grow food, a drought situation can pose major problems for agriculture. Hence, farmers often face extreme poverty in drought-prone areas. Efficient water use techniques are very important in the face of climate change.  Irrigated agriculture is placing increasing pressure on finite freshwater resources, especially in developing countries where water extraction is often unregulated, unpriced and even subsidised. To shift to a more sustainable use of water in agriculture without harming the food security and livelihoods of hundreds of millions of smallholders, substantial improvement in water use efficiency is required. Here are some ways to do it.  **1. Use of water efficient irrigation system**  Drip and sprinkler irrigation systems are the most water efficient irrigation systems. They deliver water directly to a plant’s roots, reducing the evaporation that happens with spray watering systems. Timers can be used to schedule watering during the cooler parts of the day which further reduces water loss. Many places in Maharashtra, Haryana, Meghalaya and Rajasthan where water is scarce, drip irrigation is efficient. Properly installed drip irrigation can save up to 80 percent more water than conventional irrigation and can even contribute to increased crop yields.  Screen%2BShot%2B2015-10-20%2Bat%2B1  **Fig 6.1 Drip Irrigation System**  **2. Watershed development**  Man and his environment are interdependent. The changes in the environment directly affect the lives of the people depending on it. Environmental degradation can be tackled effectively through holistic development of the watershed. Watershed provides a natural geo-hydrological unit for planning any developmental initiative. Watershed development can:   * mitigate the adverse effects of drought on crops and livestock * control desertification * encourage restoration of ecological balance * promote economic development of village community   Related image  **Fig 6.2 Watershed**  **3. Irrigation scheduling**  Modern water management is not just about how water is delivered but also when, how frequent and in what quantity. To avoid under- or over-watering crops, farmers should carefully monitor the weather, the moisture of both soil and plant and adapt their irrigation schedule to the current conditions. Some farmers water at night to slow down evaporation and to let the water seep down into the soil and replenish the water table.  Image result for Irrigation scheduling in agriculture  **Fig 6.3 Irrigation scheduling**  **4. Drought-tolerant crops**  The drought-prone areas where water scarcity is a permanent problem, growing less water intensive crops like jowar, bajra, ragi and other millets, pulses and lentils, vanilla, black pepper etc can give very good returns with less water requirement. With the advances in biotechnology, many crop cultivars with less water requirement have been released.  Screen%2BShot%2B2015-10-20%2Bat%2B1  **Fig 6.4 Drought-tolerant crops**  **5. Dry farming**  Farming techniques like the use of mulches, residue management in crop fields, etc are helpful in retaining soil moisture for crop production. Farmers who are practising dry farming, don’t use irrigation and depend on the available soil moisture to produce crops during dry season. Special tilling practices and careful attention to microclimates are essential. Dry farming tends to enhance flavours but produces lower yields than irrigated crops.  https://cuesa.org/sites/default/files/field/image/dirty_girl_dry_farm_tomato.jpg  **Fig 6.5 Dry farming**  **6. Rotational grazing**  Rotational grazing is a method in which animals are shifted between fields to promote pasture regrowth. Good grazing management increases the field’s water absorption and decreases water runoff, making pastures more drought-resistant. Increased soil organic matter and better forage cover are also water-saving benefits of rotational grazing.  Image result for Rotational grazing  **Fig 6.6 Rotational grazing**  **7. Mulch and compost**  Compost, or decomposed organic matter used as fertiliser has been found to enhance water-holding capacity and improve soil structure. Mulch is a material spread on top of the soil to retain soil moisture. Mulch can be made from organic materials like wood chips or straw which can break down into compost. It further enhances the soil’s water-holding capacity. Compost and mulch can help to retain more water in the soil during dry season. Farmers may also use black plastic mulch as soil cover to suppress weeds and reduce evaporation.  Image result for Mulch and compost  **Fig 6.7 Compost and Mulch**  **8. Cover crops**  Cover crops are meant to protect soil that would otherwise go bare. They reduce weeds and also increase soil fertility and organic matter. They also help to prevent erosion and compaction. Cover crops allow water to penetrate the soil more easily which improves its water-holding capacity. Certain research have found that fields planted with cover crops were 11 to 14 percent more productive than conventional fields during years of drought. Generally, cover crops are selected based on season and agro-climatic condition. In India, most frequently used cover crops are lentils, field peas, berseem, cowpea, soybean etc. It helps in less crop-weed competition.  Related image  **Fig 6.8 Cover crops**  **9. Conservation tillage**  Conservation tillage uses specialised ploughs or other equipment that partially till the soil but leave at least 30 percent of vegetative crop residue on the surface. Like the use of cover crops, such practices help increase water absorption and reduce evaporation, erosion, and compaction. Reduced tillage or no tillage is a very good option for environmental safety as it facilitates no burning of crop residue and enriches soil fertility by decomposing it.  Related image  **Fig 6.9 Conservation tillage**  **10. Going organic**  Recent research found that corn grown in organic fields had 30 percent greater yield than conventional fields in years of drought. In addition to keeping many of the more toxic pesticides out of waterways, organic methods help retain soil moisture. Healthy soil that is rich in organic matter and microbial life serves as a sponge that absorbs and retains moisture for plants. The trial also found that organic fields can recharge groundwater supplies up to 20 percent.  Related image  **Fig 6.10 Going organic**  **CHAPTER 7**  **SYSTEM TESING**  Testing is the process of detection errors. Testing performs a very quality role for assurance and for ensuring the ability of software. The results of testing are used latter on during maintenance also.  **6.1** **TESTING OBJECTIVES:**  The main objective of testing is to uncover a host error, systematically and the minimum effort and time starting formally, we can say   1. Test is the process executing a program with the intent of finding an error. 2. A successful test is one that uncovers as ad yet undiscovered error. 3. A good test case is one that has a high probability of finding errors, if it exits.   **6.1.1 WHITE BOX TESTING:**   * This is unit testing method where is unit will be taken at a time and tested thoroughly at a statement level to find the maximum level errors. * We have tested step wise every piece of code ,taking care that every statement in the code * Is executed at least once. * The white box testing is also called glass box testing.   **6.1.2 BLACK BOX TESTING:**  This testing method a models a single unit and checks the unit at interface and communication with other models rather getting g into details levels. Here the model will be treated as a black box that take input and generates the output. Output of given input combinations are forwarded to other models.  **6.1.3 UNIT TESTING:**  Unit testing focuses verification effort on the smallest unit of software that is the model using the detailed design and the processes specifications testing is done to uncover errors within the boundary of the model all model must be successful in the unit test before the start of the integration testing.  In our project s unit testing involves checking each future specified in the component a component performs only small part of the functionality on the system and relies on cooperating with other part of the system.  **6.1.4 INTEGRATION TESTING:**  In this project integrating all the modules forms the main system when integrating all the modules we have checked whether the integration effects working of any of the services by giving different combinations of inputs with which the services run perfectly before integration.  **6.1.5 FUNCTIONAL TESTING:**  In this test cases are decided based on specification of the system. The software or the module to be tested is treated as a black box testing and hence this also called black box testing.    **CHAPTER 8**  **DOCUMENTATION**   * 1. **HOW TO OPERATE THE SYSTEM:**  1. Install the R studio. After installing successfully, dump the sample code into the R studio by clicking File -> New -> R script and save the file with .csv extension. 2. Before running the program check whether the input file (the file which contains the reviews) is in the current working directory or not. 3. Select all the code and click on run to execute all at a while. For step by step execution use, ctrl +Enter. 4. The plots and graphs will appear in the plot area. Zoom the plot for better visualization.   **CHAPTER 9**  **SCOPE FOR FUTURE DEVELOPMENT**  Using this project, we have estimated production of some crops like Paddy, Maize, Groundnut, Sugarcane, Greengram for all the districts of Andhra Pradesh. We have used data mining tool Rstudio for Prediction of crops. We considered factors like Area, Production and Rainfall in our project. We can further explore the factors like soil, organic, and inorganic crops, use of fertilizers and pesticides. In future, we can implement this technique and can estimate for other crops. Similarly, We can use this project for predicting production of crops of all other states of India. We can even use Machine learning in this project for future implementation. By including IoT, we can encourage E-farming. We can use Cloud to store and retrieve data. When the data size which we are going to handle reaches huge sizes ,then by using big data tools like Hadoop, spark and sandbox’s can be used for more effective analysis. Using more new advanced algorithms for increased accuracy in predictions.  **CHAPTER 10**  **CONCLUSION**  As humans, we cannot predict the future production. Taking this problem into consideration we have found a solution Agriculture prediction for enhancing crop yield using Data mining tool R.  According to estimates of economic survey of A.P , 2016-17 , the food grains production was 156.85 lakh tones , and all seeds production was 24.62 lakh tones ,an increase from the previous year of 9.09% and 12.9% respectively. In the food grains segment –paddy, bajra , maize ,ragi and pulses production increased a little, where as jowar , other millets and wheat fell. In the oil seed segment , the production of groundnut and castor saw a drop while sesamum production improved.  During 2016-17,A.P had a 28% of rainfall deficit than normal .Further in the state, the total irrigated area to grow area is 50.38%. Apparently,the cro[ps during kharif season have sufficient water while farming in rabi season has suffered due to rainfall deficit, and lower level of water in reservoirs since Dec1 ,2016.  The consequences of 14.03% growth in the agriculture and allied sectors have to reflect on the upward economic mobility of the approximate 62% of the state’s population dependent on these sectors.  The positive signs would be an increase in the farm incomes, the economic status of the dependent families , and regular debt repayment by the farmers. With the government spending large budget on the farm loan waiver , free electricity supply, fertilizer subsidy, availability of credit etc., It is important to track the farm-level incomes to target and support the most vulnerable sectors involved in the agricultural and allied sectors  **REFERENCES**  [1] Grady Booch, James Rumbaugh, Ivar Jacobson : The Unified Modeling Language User Guide, Pearson Education.Rob Pandey, Pauline Wilcox  [2] Software Engineering, A Precise approach, Pankaj Jalote, Wiley.  [3] [Aggarwal et al., 2008](https://www.sciencedirect.com/science/article/pii/S0308521X16302311) P.K. Aggarwal, K.B. Hebbar, M. Venugopal, S. Rani, A. Biswal, S.P. Wani**Global theme on agroecosystems quantification of yield gaps in rain-fed rice, wheat, cotton and mustard in India** Global Theme on Agroecosystems Report No. 43. Andhra Pradesh, India (2008)  [5] [Alemaw and Simalenga, 2015](https://www.sciencedirect.com/science/article/pii/S0308521X16302311)B.F. Alemaw, T. Simalenga**Climate change impacts and adaptation in rainfed farming systems: a modeling framework for scaling-out climate smart agriculture in sub-Saharan Africa** Am. J. Clim. Chang., 4 (2015), pp. 313-329 [CrossRef](https://doi.org/10.4236/ajcc.2015.44025)[View Record in Scopus](https://www.scopus.com/inward/record.url?eid=2-s2.0-85018059469&partnerID=10&rel=R3.0.0)  [6] [Auffhammer et al., 2012](https://www.sciencedirect.com/science/article/pii/S0308521X16302311)M. Auffhammer, V. Ramanathan, J.R. Vincent**Climate change, the monsoon, and rice yield in India** Clim. Chang., 111 (2012), pp. 411-424, [10.1007/s10584-011-0208-4](https://doi.org/10.1007/s10584-011-0208-4) [CrossRef](https://doi.org/10.1007/s10584-011-0208-4)[View Record in Scopus](https://www.scopus.com/inward/record.url?eid=2-s2.0-84856746611&partnerID=10&rel=R3.0.0)  [7] [Behera et al., 2015](https://www.sciencedirect.com/science/article/pii/S0308521X16302311)R.N. Behera, D.K. Nayak, P. Andersen, I.E. Måren**From jhum to broom: agricultural land-use change and food security implications on the Meghalaya Plateau, India** Ambio (2015), pp. 63-77, [10.1007/s13280-015-0691-3](https://doi.org/10.1007/s13280-015-0691-3) [CrossRef](https://doi.org/10.1007/978-81-322-2310-8_4) .     |  | | --- | |  | |
|  |