www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 01, Issue 03, Dec. 2021, pp: 31-40

e-ISSN : xxxx-xxxx
Impact Factor : xxxx

WEATHER BASED CROP YIELD ANALYSIS AND PREDICTION USING SUITABLE ML ALGORITHMS

Varu Singh[*1], Archana Kangole[*2]

[*1]Student, Computer Science and Engineering Department, RTC Institute of Technology, Ranchi

[*2]Associate Professor, Computer Science and Engineering Department, RTC Institute of Technology, Ranchi.

ABSTRACT

Agriculture is the most prominent sector of Indian Economy, where more than 50% of population depends on agriculture. Most of the Indians are dependent on agriculture either in a direct or an indirect manner, some people are directly attached with the farming and some other people are involved in doing other businesses related to it. Fluctuations occur on the price of the paddy because of the low crop yield, consequently the common man and farmers are bearing the losses. For increasing supply of food and raw materials at non-inflationary prices we introduce crop yield prediction based on weather forecasting as most of our crops depend on weather conditions. Machine learning techniques can be used to improve prediction of weather-based crop yield prediction. The farmers will be advised for the suitable crop by predicting (estimating) the weather conditions. The type of crop and the yield are interdependent. In this scenario Rain, Temperature data sets etc. will be considered for analysis. The Machine learning algorithms and regression-based algorithms will be applied based on the requirement. The suitable user-friendly GUI will be designed and developed for this application to help the farmers and agriculture department.

KEYWORDS: Weather forecasting, Yield prediction, regression-based algorithms, GUI.

I. INTRODUCTION

Crop Yield Prediction Using Machine Learning is a Dashboard Application where a user canupload a dataset into the tool, preprocess the data with the given measures of central tendencies (mean, median, and mode) plot the graphs to visualize data, build a model to predict the results and to provide analytics. A Data Scientist directly interacts with the front-end module where the tool contains all the options for the user. Front end tool interacts with internal r-shiny server to run the application. This application works reactively where it responds to users inputsand displays the output quickly. This application can also be hosted on a remote r-shiny server toprovide it as a web- application

Purpose of system:

Crop Yield Prediction is the prediction of the yield of the given agricultural area for the given location using science and technology. This includes temperature, rain, wind speed, humidity, soil fertility etc. In addition to prediction, crop yield prediction includes predictions of changes on the earth's climate. The application has a Pre-Process tab which ensures that the data is clean and ready to be analyzed. It is the most important phase before predicting any attribute's near future. The application is open to everyone. Anyone can access the application. All the user needs to do is just upload the dataset which contains appropriate data. If the data is appropriate, then the output will be appropriate. The algorithm is efficient enough to produce the goods.

Existing System

The existing crop yield prediction applications predict the yield of any state or region but not to the yield of a person or a farmer in specific of a particular farmer. Also the predicted output may or may not be correct because it takes the conditions of the atmosphere of the whole locality. The application may or may not have the option to preprocess the data. The application must be able to ensure that the data is clean and the predicted output is appropriate.

Proposed System

Through this application, the output is predicted for any farmer's dataset or any region's dataset. Suppose one wants to know the volume of yield for a particular season next year. This application provides that facility. Also the predicted output is almost close to the accurate value. Any attribute related to yield can be predicted with good accuracy. If not the exact and accurate value, the output will be closer to the original value.

Yield Prediction

Yield Prediction is the approach to anticipate the yield of the yields utilizing various boundaries like precipitation, temperature which will go under season, region and other environmental **c**onditions and boundaries. Since the creation of new imaginative innovations and procedures the horticulture field is gradually corrupting. Because of these, bountiful development individuals are been focused on developing counterfeit items that is cross breed items where there prompts an undesirable life. These days, present day individuals don't have mindfulness about the development

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 01, Issue 03, Dec. 2021, pp: 31-40

e-ISSN : xxxx-xxxx
Impact Factor : xxxx

of the yields in an ideal time and at an ideal spot. In view of these developing methods the occasional climatic conditions are likewise being changed against the basic resources like soil, water and air which lead to instability of food. yield forecast is a significant agrarian issue. Every single rancher is consistently attempts to know, how much yield will get from his assumption. Before, yield forecast was determined by dissecting ranchers past experience on a specific harvest. The Agricultural yield is basically relies upon climate conditions, irritations and arranging of collect activity. Exact data about history of harvest yield is an significant thing for settling on choices identified with agrarian danger the executives. An exact yield forecast model can assist ranchers with choosing what to develop and when to develop. There are various ways to deal with crop yield expectation.

II. METHODOLOGY

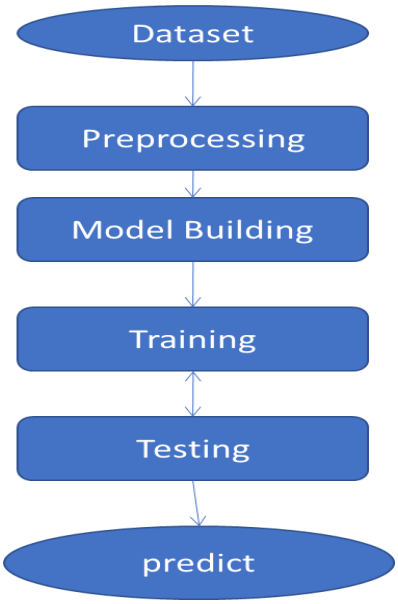

Figure 1: Data Flow Diagram

Predictions Techniques:

The given dataset can make predictions using different machine leaning algorithms in a single application. The available techniques are:

- Multi linear Regression
- Random Forest
- Neural Network
- Subspace Clustering
- K- means Clustering

SYSTEM ANALYSIS:

Feasibility Study

The feasibility of the assignment is analyzed in this segment and enterprise idea is placed forth with a very preferred plan for the venture and a few fee estimates. During device analysis the feasibility have a look at of the proposed machine is to be carried out. For feasibility analysis, major key factor the machine is essential.

Three types of feasibility analysis are

Economical Feasibility

Technical Feasibility

Social Feasibility

- **Economical Feasibility**

This feasiability take a look to check the financial impact that the gadget could have at the organization. The expenditures must be justified. Thus the developed gadget as nicely inside the price range and this was executed because maximum of the technologies used are freely to be had.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 01, Issue 03, Dec. 2021, pp: 31-40

e-ISSN : xxxx-xxxx
Impact Factor : xxxx

- **Technical Feasibility**

This have look at is carried out to test the technical feasibility, that is, the technical requirements of the gadget. Any machine developed have to not have a high demand on the to be had technical resources. This will lead to excessive needs on the available technical resources. This will lead to high demands being located on the client. The developed machine must have a modest requirement, as best minimum or null changes are required for implementing this machine.

- **Social Feasibility**

The factor of observe is to check the extent of recognition of the machine by means of the consumer. This consists of the technique of education the user to use the gadget efficiently. The

person have to no longer experience threatened by the device, instead must receive it as a necessity. The degree of acceptance by the users solely relies upon on the methods which can be hired to educate the user approximately the system and to make him acquainted with it. His level of confidence have to be raised so that he's also able to make a few constructive criticism, that's welcomed, as he's the final consumer of the gadget.

III. MODELING AND ANALYSIS

Architectural Design

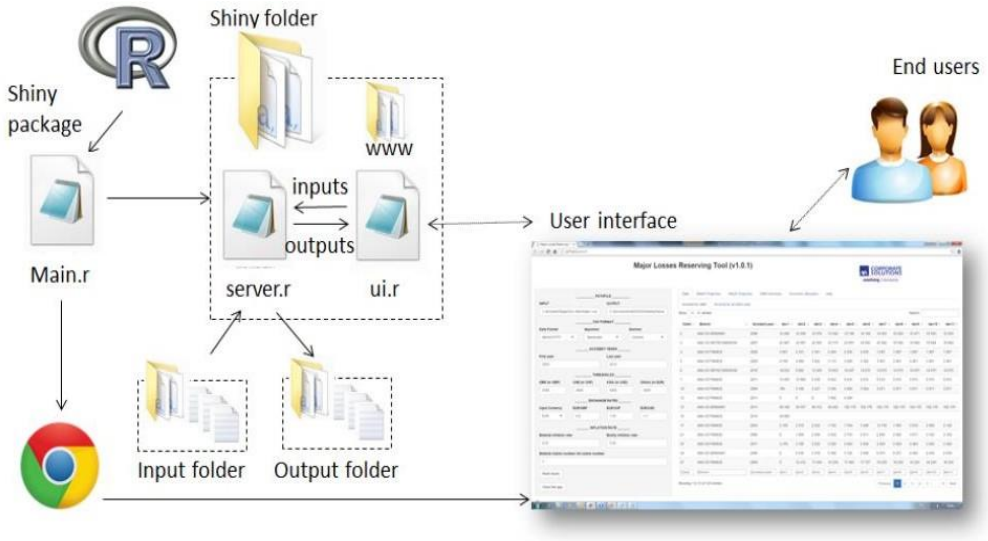

Figure 2: Architecture Diagram

UML Diagrams

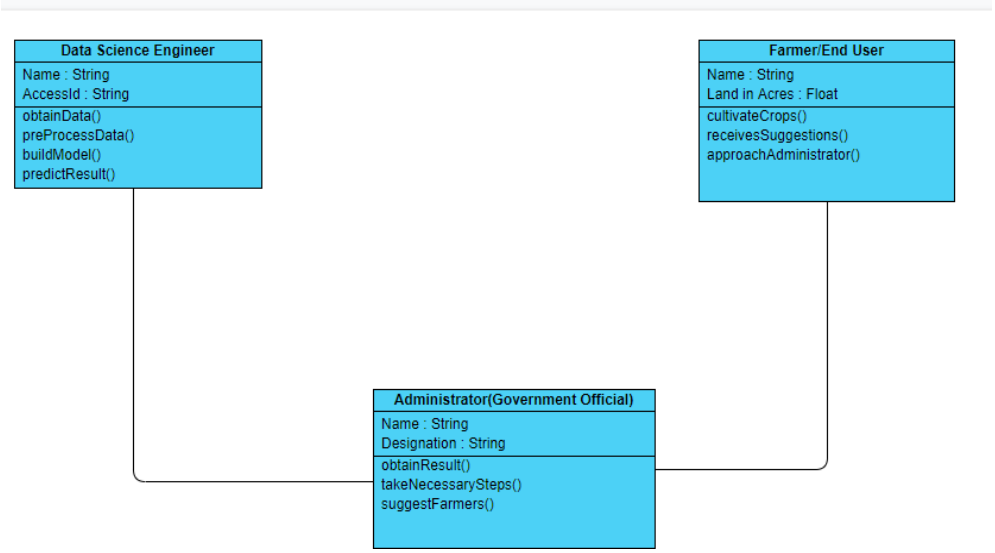

Figure 3: Class Diagram

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 01, Issue 03, Dec. 2021, pp: 31-40

e-ISSN : xxxx-xxxx
Impact Factor : xxxx

The above diagram shows the class diagram. The class diagram has various classes used in our system and the relationship among them. Classes are represented in the form of rectangles. The top most portion of the rectangle is the class name, the middle portion has the attributes and their access specifies and the bottom most portion consists of the methods.

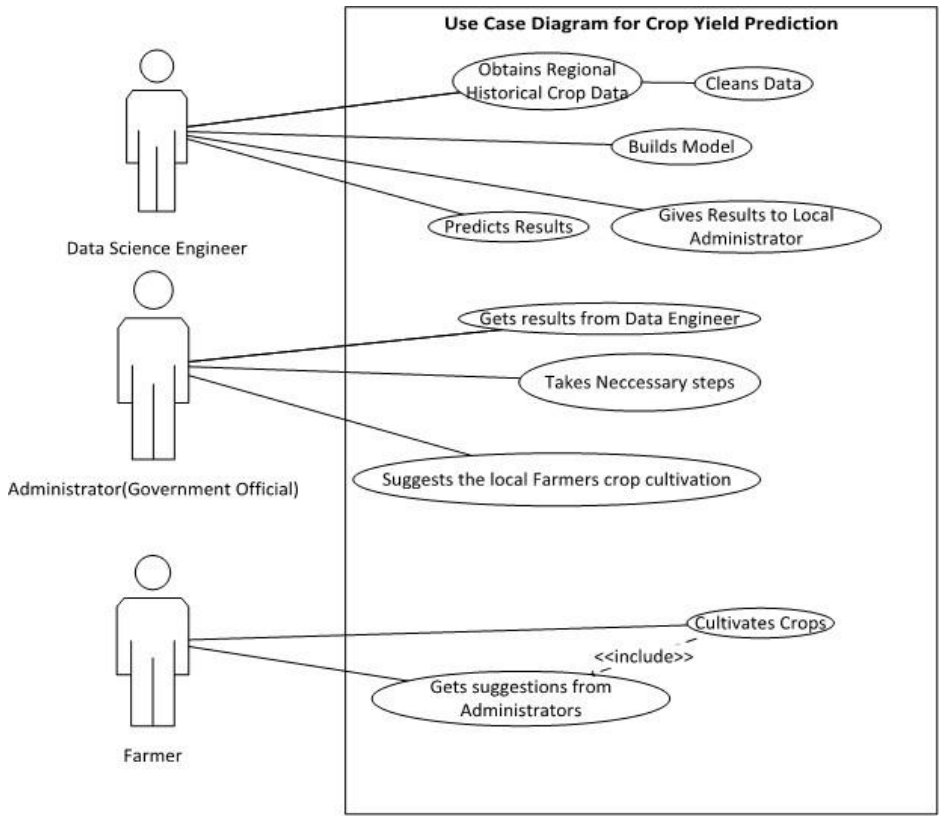

Figure 4: Use Case Diagram

The above diagram shows the use case diagram of the project. The actors are user and server. The actions performed by the actors are being depicted in the oval shaped diagrams. The lines connecting actors and actions mean that the particular actor is performing those actions.

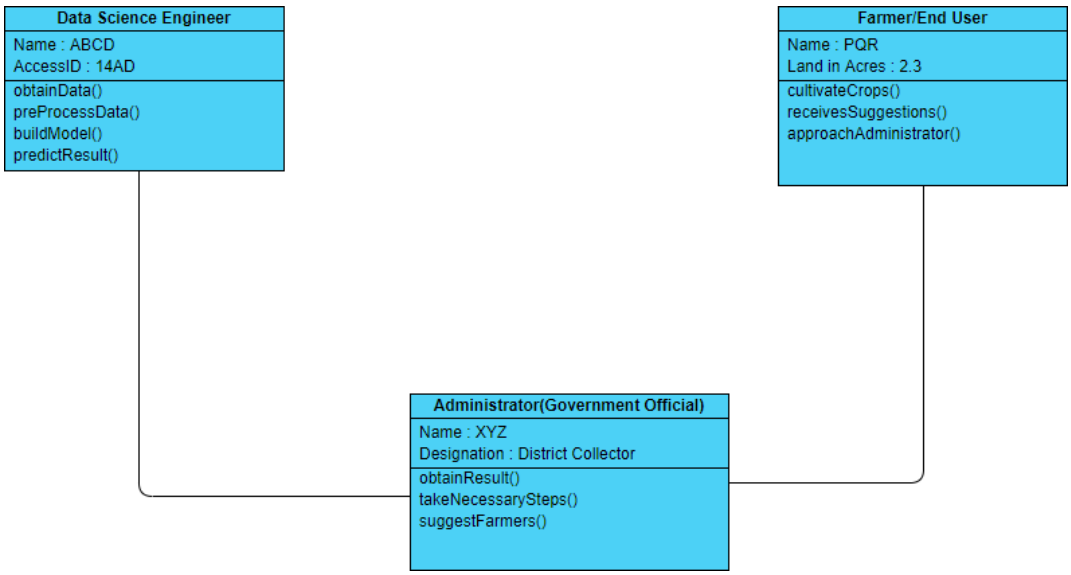

Figure 5: Object Diagram

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 01, Issue 03, Dec. 2021, pp: 31-40

e-ISSN : xxxx-xxxx
Impact Factor : xxxx

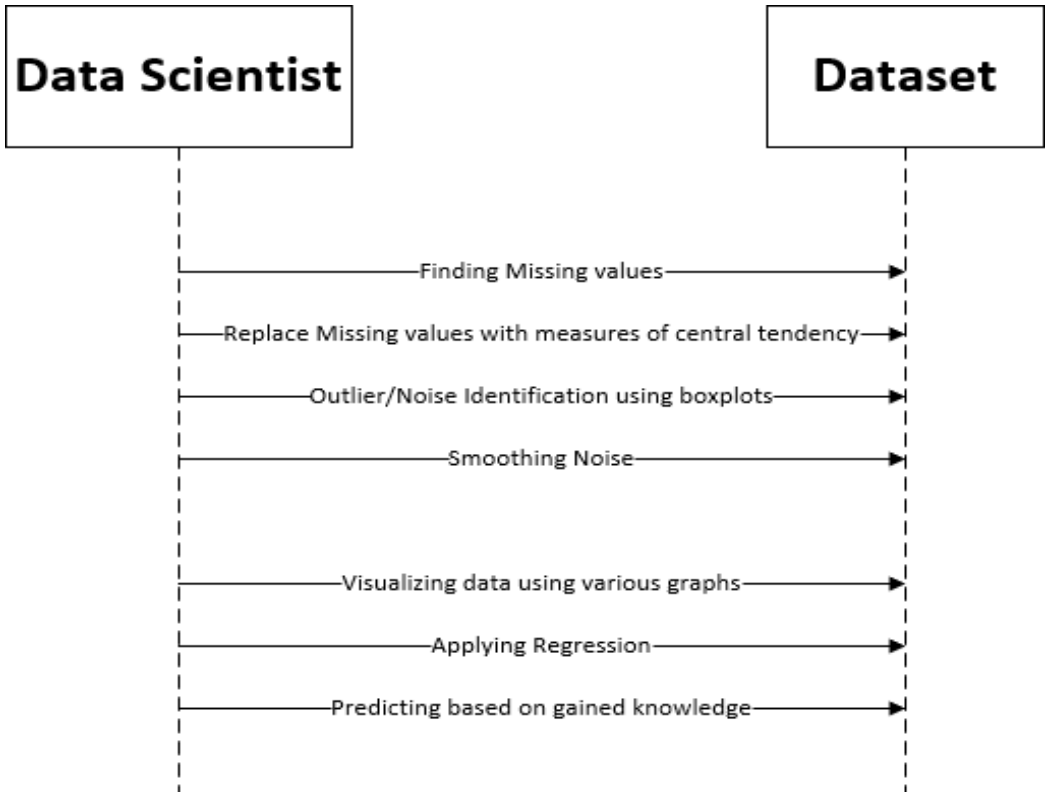

Figure 6: Sequence Diagram

Sequence Diagram shows the sequence of events. The farmer requests the government admin for suggestions which the government official gets from the data scientist. The data scientists upload the data set and preprocess it, plot graphs and predicts the output. Data Scientist forwards the data results to government admin.

SYSTEM IMPLEMENTATION

The implementation level of any assignment is a true display of the defining moments that make a project a success or a failure. The implementation level is defined as the system or device changes being hooked up and made operational in a manufacturing environment. The phase is initiated after the machine has been examined and accepted through the user. This section continues until the machine is running in production in accordance with the defined person requirements

Algorithm

Algorithm Crop Yield Prediction

{

1. Upload the dataset by browsing already downloaded dataset from the current system.
2. After the dataset gets uploaded the user can either view or edit the records.
3. The next step will be preprocessing.
4. Missing values and NANs are replaced by any value of mean, median or mode.
5. These values are replaced for the selected attributes.
6. If there are no missing values or NANs, then, the dataset remains the same after preprocessing.
7. After preprocessing, graphs are plotted for several attributes.
8. Graphs are plotted to analyze the nature of the attributes.
9. For a graph to be plotted, select at least two attributes.
10. After plotting the graph, click on the Prediction tab.
11. In the prediction tab, select the attribute for which the output should be predicted.
12. The output will be predicted for the selected attribute.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 01, Issue 03, Dec. 2021, pp: 31-40

e-ISSN : xxxx-xxxx
Impact Factor : xxxx

SYSTEM TESTING

The purpose of checking out is to find out errors. Testing is the method of attempting to find out every achievable fault or weak point in a work product.

Table 1: Test Case For Validation Of Tabs

TEST SCENARIO	CASE ID	TEST DATA	EXPECTED OUTPUT	ACTUAL OUTPUT	TEST PASS/FAIL
VALIDATION OF TABS	TC_UT_01	i) BROWSE DATA SET ON THE SYSTEM ii.) BROWSE DATA SET ON THE SYSTEM	DATA SET UPLOADED DATA SET UPLOADED	DATA SET UPLOADED INVALID FORMAT	PASS FAIL
	TC_UT_02	i)EDIT YOUR DATA	UPDATION SUCCESSFUL	UPDATION SUCCESSFUL	PASS
	TC_GT_03	DISPLAY GRAPH	GRAPH DISPLAYED	GRAPH DISPLAYED	PASS

Table 2: Test Case to Check Functionality

TEST SCANARIO	TEST CASE ID	TEST DATA	EXPECTED OUTPUT	ACTUAL OUTPUT	TEST PASS/FAIL
VALIDATIO N OF TABS	TC_UT_0 1	i.)BROWSE DATA SET ON THE SYSTEM ii.)BROWSE DATA SET ON THE SYSTEM	DATA SET UPLOADED DATA SET UPLOADED	DATA SET UPLOADED DATA SET UPLOADED	PASS PASS
	TC_UT_0 2	i)EDIT YOUR DATA	UPDATION SUCCESSFULLY	UPDATION SUCCESSFULLY	PASS

IV. RESULT AND DISCUSSION

Upload:

In the upload tab user can do the following operations like:

Browse:

In the browsing option user can upload any data set he wants from local computer.After uploading the file User has uploaded will be displayed

View Your Data:

In the view your data option user can view the data he/she upload in the separate screen, in this he will get a Clear picture of what kind of data he has uploaded.

Edit Your Data:

In the edit your data tab user can edit data however he likes it, in this he/she can remove the columns or Change the data fields.

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 01, Issue 03, Dec. 2021, pp: 31-40

e-ISSN : xxxx-xxxx
Impact Factor : xxxx

Select Columns:

In the Select columns tab user can limit the number of columns as he/she wishes according to the operation he Wants to perform such as Analytics, Modeling etc.

Search:

In the search option user can be able to see any data or any particular field he view bysearching the keywords Related to that particular data.

PREPROCESS:

Select columns for preprocess:

This tab is used to select the field user wants to preprocess. Preprocess can include replacing null values with Mean, Median, Mode.

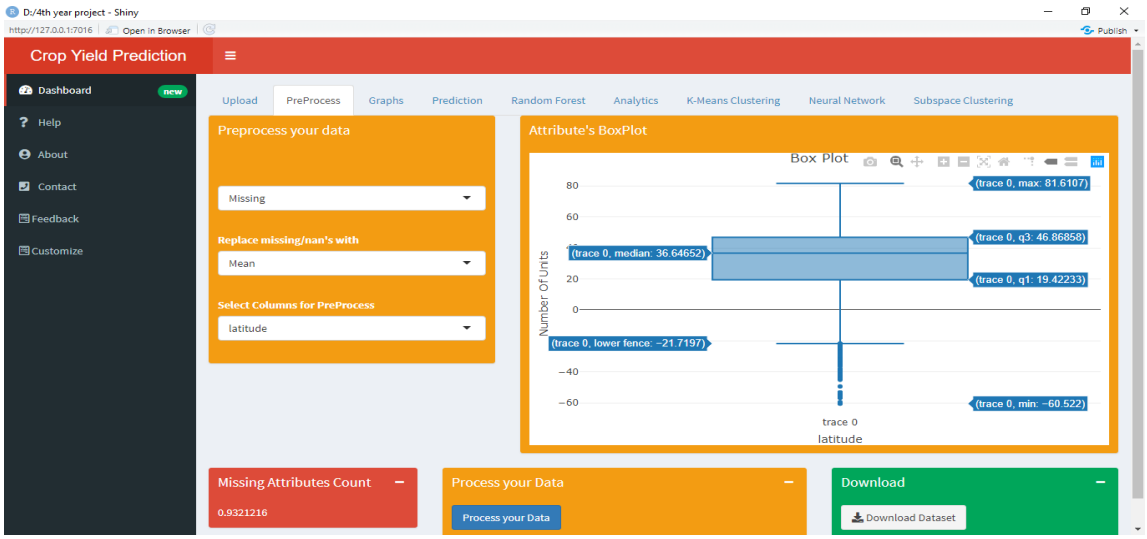

Figure 7: Displaying Box plot

Graphs:

1. Select Graph to Display:

In the Select graph to display drop down box user use can select the type of plot he wants to draw this includes:

1) Box plot. 2) Scatter plot 3) Histogram 4) Bar plot

2. Select 1st column for plot:

In this drop down box user needs to select the first field in drawing that particular graphs he has selected.

3. Select 2nd column for plot:

In this drop down box user needs to select the first field in drawing that particular graphs he has selected.

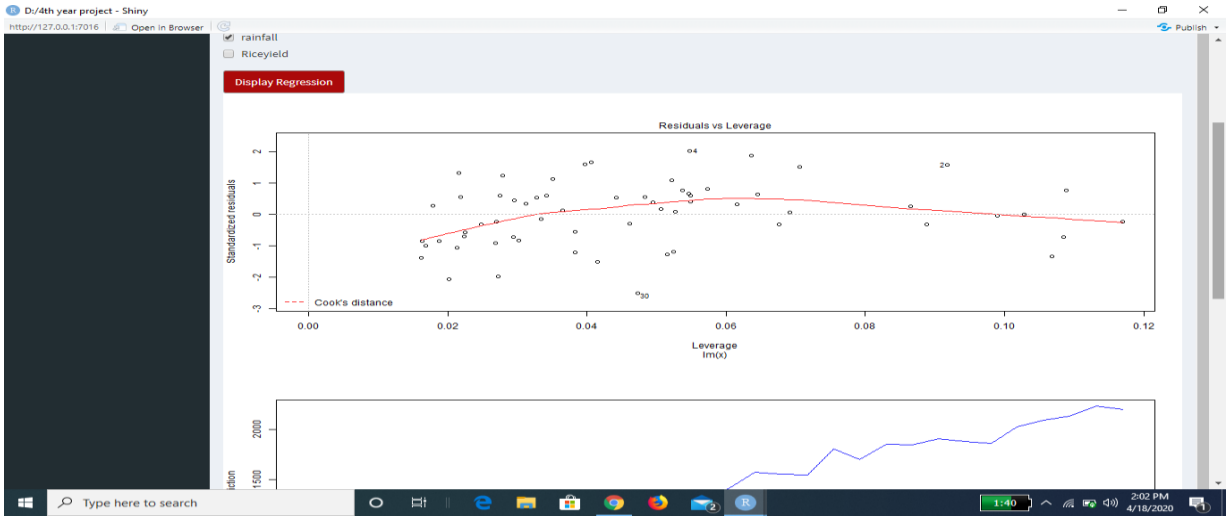

Figure 8: Scatter plot for Year and Yield

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 01, Issue 03, Dec. 2021, pp: 31-40

e-ISSN : xxxx-xxxx
Impact Factor : xxxx

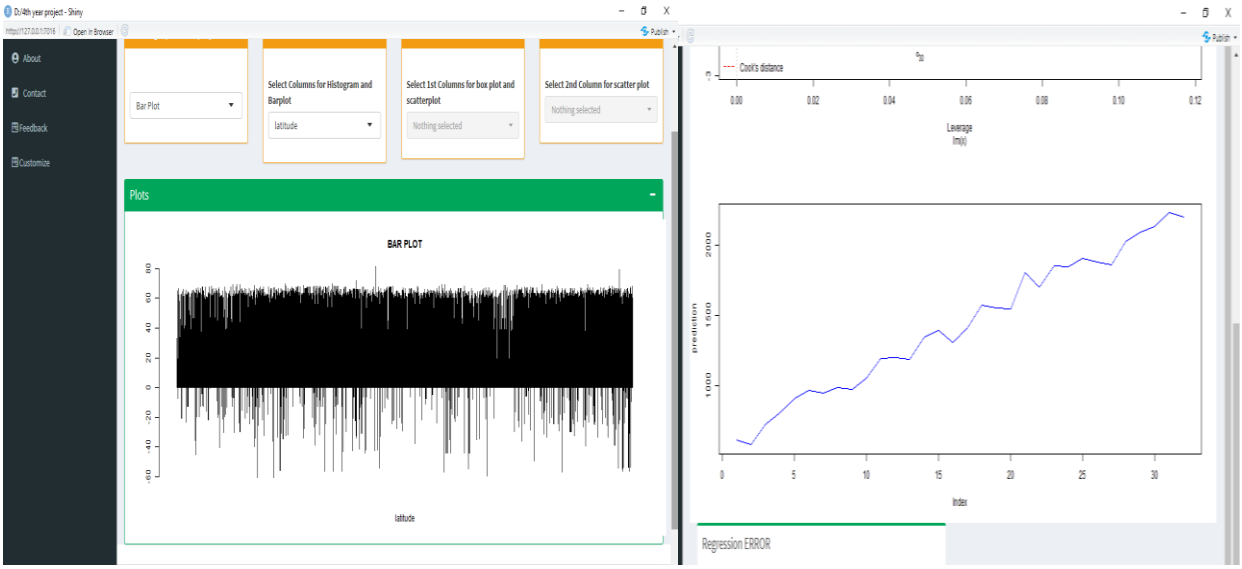

Figure 9: Display Graph

Figure 10: Scatter Plot

Predictions:

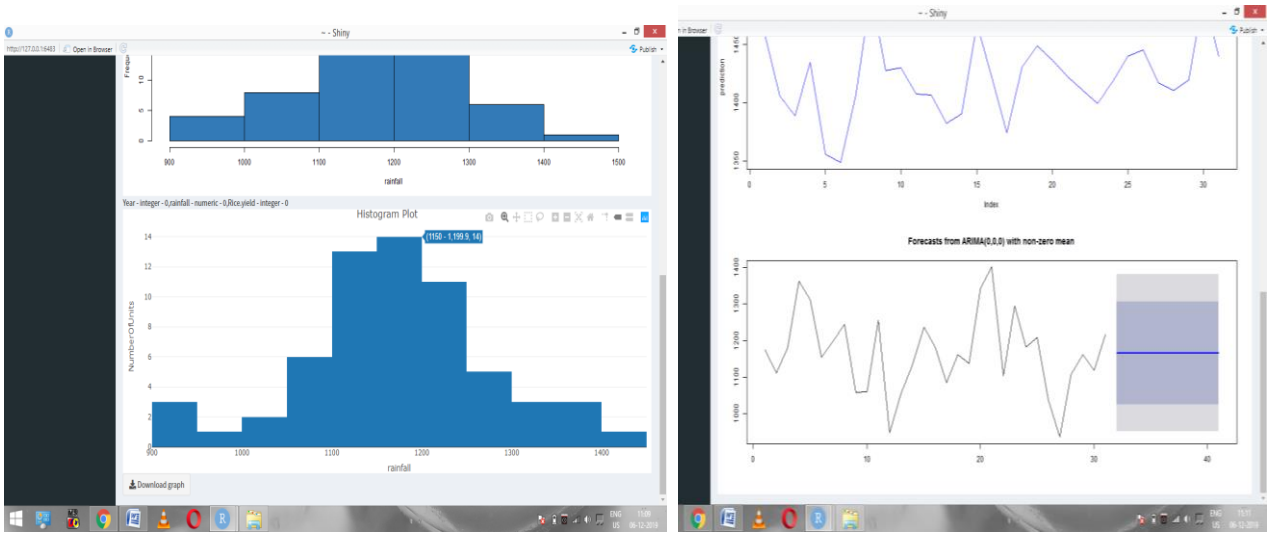

Figure 11: Histogram

Figure 12: Forecast for Rainfall and Yield

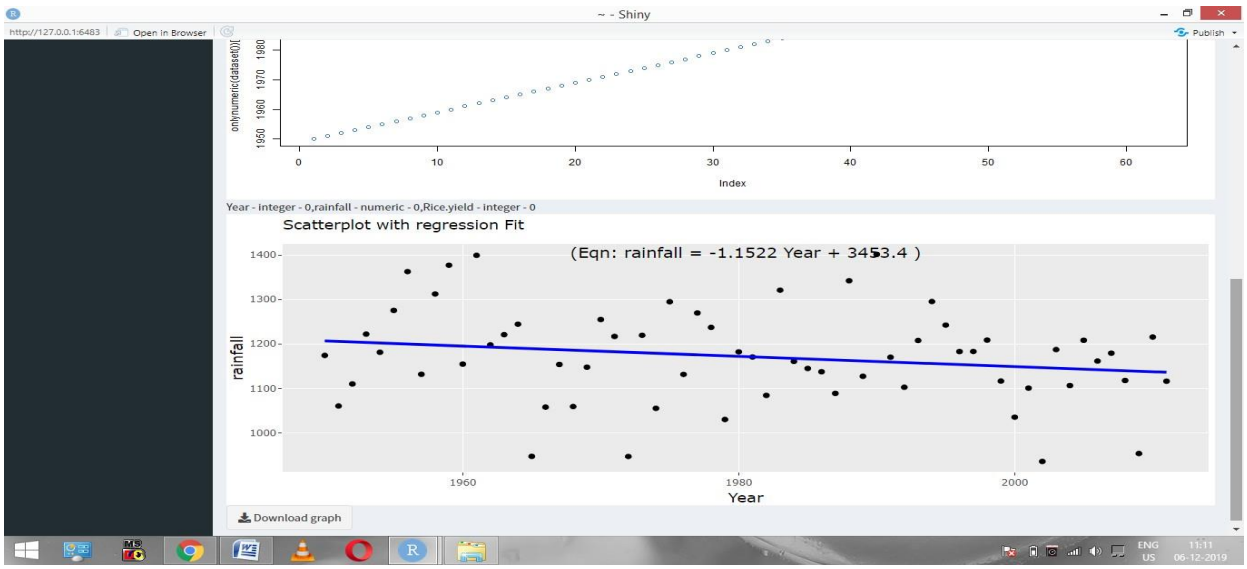

Figure 13: Scatter plot with Regression Fit

www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
Vol. 01, Issue 03, Dec. 2021, pp: 31-40

e-ISSN : xxxx-xxxx
Impact Factor : xxxx

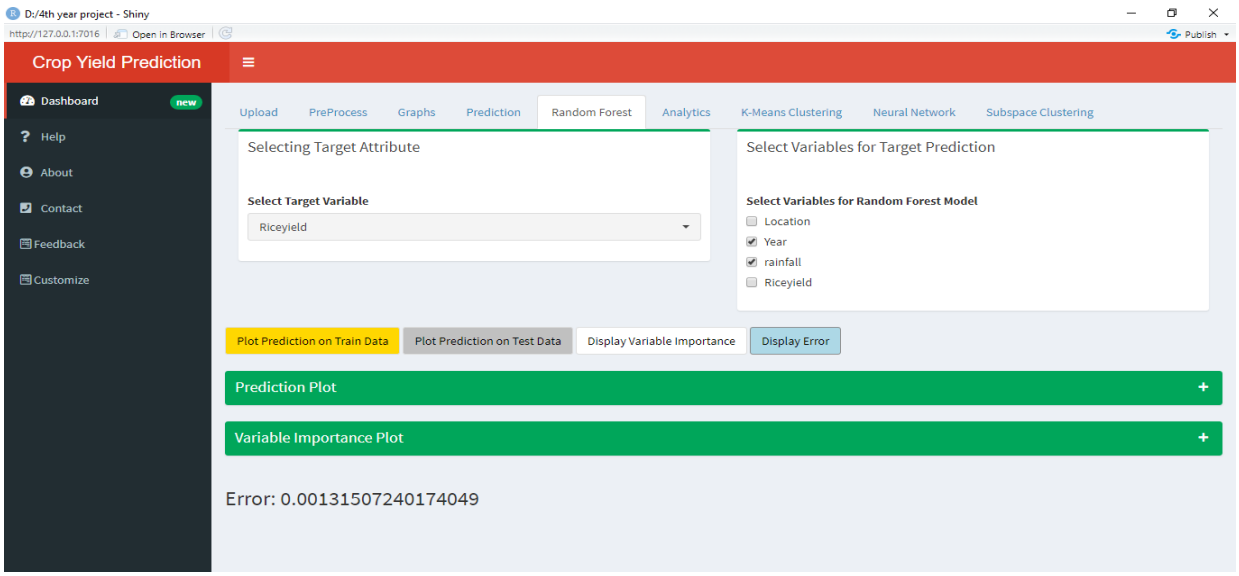

Figure 14: Random Forest

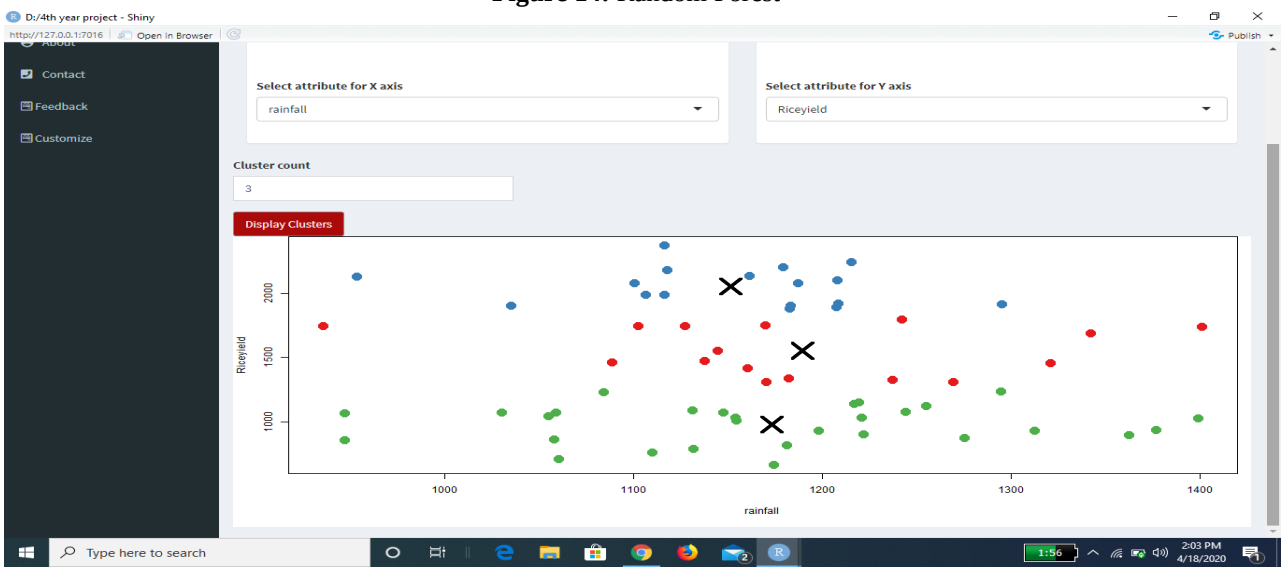

Figure 15: K-means cluster (rainfall v/s rice yield)

V. CONCLUSION

This application will be of a great use to the administrators and will be of a great help to the farming community of India. This application can be used by administrators to get an estimate of yield which can help them in taking appropriate actions like procuring/importing food grains at times of crisis or making arrangements for proper export during surplus yield. The farmers and can understand the kind of crop to be grown as the dataset is region and crop centric.

VI. REFERENCES

[1] P.Priya, U.Muthaiah M.Balamurugan.Predicting yield of the crop using machine learning algorithm. International Journal of Engineering Science Research Technology.

[2] J.Jeong, J.Resop, N.Mueller and team.Random forests for global and regional crop yield prediction.PLoS ONE Journal.

[3] Narayanan Balkrishnan and Dr. Govindarajan Muthukumarasamy.Crop production Ensemble Machine Learning model for prediction. International Journal of Computer Science and Software Engineering (IJCSSE).

[4] S.Veenadhari, Dr. Bharat Misra, Dr. CD Singh.Machine learning approach for forecasting crop yield based on climatic parameters. International Conference on Computer Communication and Informatics (ICCCI).

[5] Shweta K Shahane , Prajakta V Tawale.Prediction On Crop Cultivation. IInternational Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE) Volume 5, Issue 10, October 2016.

[6] D Ramesh ,B Vishnu Vardhan. Analysis Of Crop Yield Prediction Using Data Mining Techniques. IJRET: International Journal of Research in Engineering and Technology.

[7] Subhadra Mishra,Debahuti Mishra, Gour Hari Santra. Applications of Machine Learning Techniques in Agricultural Crop Production. Indian Journal of Science and Technology, Vol 9(38), DOI:10.17485/ijst/2016/v9i38/95032, October 2016.

[8] Konstantinos G. Liakos,Patrizia Busato,Dimitrios Moshou, Simon Pearson ID,Dionysis Bochtis. Machine Learning in Agriculture. Lincoln Institute for Agri-food Technology (LIAT), University of Lincoln, Brayford Way, Brayford Pool,Lincoln LN6 7TS, UK, spearson@lincoln.ac.uk.

[9] Baisali Ghosh. A Study to Determine Yield for Crop Insurance using Precision Agriculture on an Aerial Platform. Symbiosis Institute of Geoinformatics Symbiosis International University 5th & 6th Floor, Atur Centre, Gokhale Cross Road, Model Colony, Pune – 411016.

[10] Jig Han Jeong, Jonathan P. Resop, Nathaniel D. Mueller, David H. Fleisher ,Kyungdahm Yun, Ethan E. Butler,Soo-Hyung Kim. Random Forests for Global and Regional Crop Yield Predictions. Institute on the Environment, University of Minnesota, St. Paul, MN 55108, United States of America.

[11] Ecochem Online. (2009). Soil Health and Crop yields. Last modified January 28th 2009.Retrieved on March 4th 2009 from http://ecochem.com/healthy_soil.html