International Journal of Engineering and Science Invention (IJESI)
ISSN (Online): 2319-6734, ISSN (Print): 2319-6726
www.ijesi.org ||Volume 13 Issue 4 April 2024 || PP 10-39

"Crop Raising by Predicting and Analysis of CropsUsing ML Techniques"

Darshan Morkar
(BE Computer Science and Engineering)
S. G. BALEKUNDRI INSTITUTE OF TECHNOLOGY,BELAGAVI.

Abstract
Yield prediction is a very important issue in agricultural. Any farmer is interested in knowing how much yield he is about to expect. Analyze the various related attributes like location, pH value from which alkalinity of the soil is determined. Along with it, percentage of nutrients like Nitrogen (N), Phosphorous (P), and Potassium (K), type of soil, nutrient value of the soil, rainfall and soil composition can be determined. The attributes of data is trained and analyzed using K-nearest neighbor algorithm to get the precise prediction of the crop. The system model is precise in predicting crop yield and deliver the farmer with proper recommendations about required fertilizer ratio based on atmosphere and soil parameters of the land which enhance to increase the crop yield and farmer revenue.

Keywords: *Agriculture, KNN, prediction*

Date of Submission: 02-04-2024 Date of Acceptance: 12-04-2024

I. INTRODUCTION

Agriculture is the backbone of the Indian economy. In India, agricultural yield primarily depends on weather conditions. This system's future crop productivity and an analysis are made in order to help the farmers to maximize the crop production. In the past farmers used to predict their yield from previous year yield experiences. Thus, for this kind of dataanalytics in crop prediction, there are different techniques or algorithms, and with the help of those algorithms the prediction of crop yield can be done.

Weather forecasting using ML (Machine Learning) is considered for collecting historical weather data from various weather stations to forecast weather conditions. Rainfall prediction using ML Techniques is taken into account to get the knowledge in predicting the weather for crop prediction. Prediction of Crop yield using ML to know about the prediction crop with respect to atmospheric & soil parameters. Sugarcane Crop prediction Using Supervised ML is considered in order to get to predict the unique crop by applying descriptive analytics using three datasets like as Soil dataset, Rainfall dataset, and Yield dataset as a combined dataset [1].

Machine learning methods for crop yield prediction and climate change impact assessment in agriculture is proposed semi parametric variant of a deep neural network model for crop prediction and evaluate the effects of climate change. Groundnut Prediction Using ML Techniques is taken soil, environment and abiotic attributes for predicting the groundnut yield using different ML algorithms. The accuracy of the prediction was compared using RMSE (Root Mean Square Error). Algorithms for Crop Prediction deals with estimation of crop yield from precipitation and soil input. For prediction, supervised learning algorithms were used. The algorithms were compared using MSE (Mean Square Error) for finding an optimal crop prediction. Agriculture Analysis for Next Generation High Tech Farming in Data Mining discuss MLR (Machine Learning using R) method for analyzing crops and decision tree algorithm for classification of more than 350 data. ML for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks detailed the Back Propagation Network to estimate the testing data. The hidden layers of Back Propagation Network From ancient days, agriculture is considered as the main source of supply to satisfy the daily needs of human lives. It is also considered a primary occupation, and also one of the India's major industrial sectors. The farmers are ought to follow a traditional naked eye observation and yielded healthy crops without the involvement of chemicals for animals and also to their cultivation land in order to keep healthy diversity [2].

But nowadays, weather conditions are being rapidly changing against the elemental assets to deplete the food and increase the security. In meantime, the GDP (Gross Domestic Product) in agricultural sector is keep on decreasing, where in 2005 it was about 17.2%, in 2012 it was 11.1, in 2018 it was 5% and in first quarterly year of 2109- 2020 it came downto 2%. Approximately 80 percent of farmers come from rural areas, and if the revenue from crop production goes down, their lifestyle would be influenced by the farms at industry level. This makes sense to farmers in India to show some special concern towards effective and precision farming. In India there are multiple ways to rise the crop learn profit and improve the standard of the crops so as

to keep up the economic growth within the field of agriculture. So, the deployment of one of the recent advancements in technology such as, ML is one among the answer for predicting the crop with relation to atmospheric & soil parameter of the agricultural land. Since, now-a-day's climatic conditions aren't predictable like decades ago. It is changing day by day due to globalization. Hence, the farmers are facing difficulties in forecasting the weather and crops based on climate data.

In recent years the advancement of ML plays a crucial role in every field including agriculture, here the crop prediction process done with consolidating the preceding data and the present data of a particular month to prove the accuracy of climatic data. Machine learning may be a methodology of analyzing information to automatize the given model and may be a branch of AI depend on the concept that systems will study from data to form selections with minimal human intervention. There may be a logical classifier, where a naive mathematician who predicts membership opportunities for each group, such as the possibility that knowledge belongs to a specific class [3].

The proposed system analyzes the application of supervised ML approaches the class with the very best chance is taken into account as the possibly class. Here the category is nothing however the crop that get foretold for the given input parameters. Once the crop is foretold, it will facilitate the farmers to predict the affordable crop for their individual land. Then, the farmers is guided with an application in mobile tend to make them to understand that what quite seeds we will tend to sow in land to induce higher yielding. Within the past preceding data, crop prediction was calculated by analyzing farmer's previous expertise on climatic condition. So, the correct data regarding history of climatic condition is a vital factor for creating selections in choosing crops. Therefore, this paper proposes a thought to predict the affordable crop for the given input parameter for the poorfarmers using machine learning. Thereby this proposed work will suggest the farmers witheffective solutions for more profitable cultivation [4].

The history of agriculture in India dates back to the Indus Valley Civilization Era. India ranks second in this sector. Agriculture and allied sectors like forestry and fisheriesaccount for 15.4 percent of the GDP with about 31 percent of the workforce. India ranks first globally with the highest net cropped area followed by US and China. Agriculture is demographically the broadest economic sector and plays a significant role in the overall socio-economic fabric of India. Due to the revolution in industrialization, the economic contribution of agriculture to India's GDP is steadily declining with the country's broad- based economic growth [5].

II. LITERATURE SURVEY

In [1], Data mining is the practice of examining and deriving purposeful information from the data. Data mining finds its application in various fields like finance, retail, medicine, agriculture etc. Data mining in agriculture is used for analyzing the various biotic and abiotic factors. Agriculture in India plays a predominant role in economy and employment. The common problem existing among the Indian farmers are they don't choose the right crop based on their soil requirements. Due to this they face a serious setback in productivity. This problem of the farmers has been addressed through precision agriculture. Precision agriculture is a modern farming technique that uses research data of soil characteristics, soil types, crop yield data collection and suggests the farmers the right crop based on their site-specific parameters. This reduces the wrong choice on a crop and increase in productivity. In this paper, this problem is solved by proposing a recommendation system through an ensemble model with majority voting technique using Random tree, CHAID, K-Nearest Neighbor and Naive Bayes as learners to recommend a crop for the site-specific parameters with high accuracy and efficiency.

In [2], Agriculture planning plays a significant role in economic growth and food security of ago-based country. Selection of crops is an important issue for agriculture planning. It depends on various parameters such as production rate, market price and government policies. Many researchers studied prediction of yield rate of crop, prediction of weather, soil classification and crop classification for agriculture planning using statistics methods or machine learning techniques. If there is more than one option to plant a crop at a time using limited land resource, then selection of crop is a puzzle. This paper proposed a method named Crop Selection Method (CSM) to solve crop selection problem, and maximize net yield rate of crop over season and subsequently achieves maximum economic growth of the country. The proposed method may improve net yield rate of crops.

In [3], Agriculture sectors in India is facing rigorous problem to maximize the crop productivity. More than 60 percent of the crop still depends on monsoon rainfall. Recent developments in Information Technology for agriculture field has become an interesting research area to predict the crop yield. The problem of yield prediction is a major problem that remains to be solved based on available data. Data mining techniques are the better choices for this purpose. Different Data Mining techniques are used and evaluated in agriculture for estimating the future year's crop production. This paper presents a brief analysis of crop yield prediction using Multiple Linear Regression (MLR) technique and Density based clustering technique for the selected region. The paper hypothesizes analysis of Explorative Data and considers the design of different types of predictive models. A data set is taken as a sample data set, and different regression techniques are tried to recognize and

examine each property. Specific regression methods discussed here are Multiple Linear, Linear, Non-Linear, Polynomial, Ridge regression and Logistic. Using this article, we obtain a comparative study of the different algorithms in data analytics. This helped in determining which algorithm is most appropriate to the proposed system.

In [4], the attributes in the dataset included the soil type, groundwater level, rainfall, water availability, temperature of one dataset and the other dataset included the potassium, phosphorus, and nitrogen values, fertilizers, soil pH and organic carbon value. The dataset was preprocessed using basic preprocessing tasks. Naive Bayes and J48 classifiers were used for the crop recommendation. The final recommendation was done using association rules based on the results obtained from the classifiers. The model was trained using 10- cross validation. The testing was done based on different metrics like the Accuracy, ROC Area, Recall, Precision, F-Measure etc.

In [5], the agriculture plays a dominant role in the growth of the country's economy. Climate and other environmental changes has become a major threat in the agriculture field. Machine learning (ML) is an essential approach for achieving practical and effective solutions for this problem. Crop Yield Prediction involves predicting yield of the crop from available historical available data like weather parameter, soil parameter and historic crop yield. This paper focus on predicting the yield of the crop based on the existing data by using Random Forest algorithm. Real data of Tami land were used for building the models and the models were tested with samples. The prediction will help to the farmer to predict the yield of the crop before cultivating onto the agriculture field. To predict the crop yield in future accurately Random Forest, a most powerful and popular supervised machine learning algorithm is used.

In [6], this paper has been prepared as an effort to reassess the research studies on the relevance of machine learning techniques in the domain of agricultural crop production. Methods/Statistical Analysis: This method is a new approach for production of agricultural crop management. Accurate and timely forecasts of crop production are necessary for important policy decisions like import-export, pricing marketing distribution etc. whichare issued by the directorate of economics and statistics. However, one has understood thatthese prior estimates are not the objective estimates as this estimate requires lots ofdescriptive assessment based on many different qualitative factors. Hence there is a requirement to develop statistically sound objective prediction of crop production. That development in computing and information storage has provided large amount of data. Findings: The problem has been to intricate knowledge from this raw data, this has led to the development of new approach and techniques such as machine learning that can be used to unite the knowledge of the data with crop yield evaluation. This research has been intended to evaluate these innovative techniques such that significant relationship can be found by their applications to the various variables present in the data base. Application/Improvement: The few techniques like artificial neural networks, Information Fuzzy Network, Decision Tree, Regression Analysis, Bayesian belief network. Time series analysis, Markov chain model, k-means clustering, k nearest neighbor, and support vector machine are applied in the domain of agriculture were presented.

In [7], Data Mining is emerging research field in crop yield analysis. Yield prediction is a very important issue in agricultural. Any farmer is interested in knowing how much yield he is about to expect. In the past, yield prediction was performed by considering farmer's experience on particular field and crop. The yield prediction is a major issue that remainsto be solved based on available data. Data mining techniques are the better choice for this purpose. Different Data Mining techniques are used and evaluated in agriculture for estimating the future year's crop production. This research proposes and implements a system to predict crop yield from previous data. This is achieved by applying association rule mining on agriculture data. This research focuses on creation of a prediction model which may be used to future prediction of crop yield. This paper presents a brief analysisof crop yield prediction using data mining technique based on association rules for the selected region i.e., district of Tamil Nadu in India. The experimental results shows thatthe proposed work efficiently predict the crop yield production.

In [8], in this paper we considered various situations of climatologically phenomena affecting local weather conditions in various parts of the world. These weather conditions have a direct effect on crop yield. Various researches have been done exploring theconnections between large-scale climatologically phenomena and crop yield. Artificial neural networks have been demonstrated to be powerful tools for modeling and prediction, to increase their effectiveness. Crop prediction methodology is used to predict the suitable crop by sensing various parameter of soil and also parameter related to atmosphere.Parameters like type of soil, PH, nitrogen, phosphate, potassium, organic carbon, calcium, magnesium, Sulphur, manganese, copper, iron, depth, temperature, rainfall, humidity. For that purpose, we are used artificial neural network (ANN).

In [9], an important issue for agricultural planning purposes is the accurate yield estimation for the numerous crops involved in the planning. Machine learning (ML) is an essential approach for achieving practical and effective solutions for this problem. Many comparisons of ML methods for yield prediction have been made, seeking for the most accurate technique. Generally, the number of evaluated crops and

techniques is too lowand does not provide enough information for agricultural planning purposes. This paper compares the predictive accuracy of ML and linear regression techniques for crop yield prediction in ten crop datasets. Multiple linear regression, M5-Prime regression trees, perceptron multilayer neural networks, support vector regression and k-nearest neighbor methods were ranked. Four accuracy metrics were used to validate the models: the root mean square error (RMS), root relative square error (RRSE), normalized mean absolute error (MAE), and correlation factor (R). Real data of an irrigation zone of Mexico were used for building the models. Models were tested with samples of two consecutive years. The results show that M5- Prime and k-nearest neighbor techniques obtain the lowest average RMSE errors (5.14 and 4.91), the lowest RRSE errors (79.46% and 79.78%), the lowest average MAE errors (18.12% and 19.42%), and the highest average correlation factors (0.41 and 0.42). Since M5-Prime achieves the largest number of crop yield models with the lowest errors, it is a very suitable tool for massive crop yield prediction in agricultural planning.

III. PROBLEM DEFINITION AND OBJECTIVES OF THEPROJECT

3.1 Problem Definition

The goal of the project is to help the farmers choose a suitable crop to grow in order to maximize the yield. The proposed system tries to overcome this existing system and make predictions by analyzing structured data. The solution we are proposing is to design a system taking into consideration the most influencing parameters to grow a crop and to geta better selection of crops which can be grown over the season. This would help reduce the difficulties faced by the farmers in selecting the crop to get high yield and thus maximize profits which in turn will reduce the suicide rates.

3.2 Objectives

The objectives of this project are:

- Predicting the crop yield by using K-nearest neighbor algorithm.
- Suggesting the crop to be yield based on past crop production.
- Suggesting the fertilizer which gives efficient yield.

IV. SYSTEM REQUIREMENTS

4.1 Hardware Requirements

- System with the standard minimum configuration.

4.2 Software Requirements

- IDE : PyCharm 2017 (Run and Debug the code)
- Database : MySQL (Store and retrieve data)
- Framework : Django (Design and develop)
- Frontend : HTML,CSS, and JavaScript
- Browser : Google Chrome, Microsoft Edge.

V. System Design

SYSTEM DESIGN

This system shows the communication between farmer and admin with the web application. The following are the two sections of the system design; one section elaborates the steps with respect to farmer (user) and the other section with respect to admin.

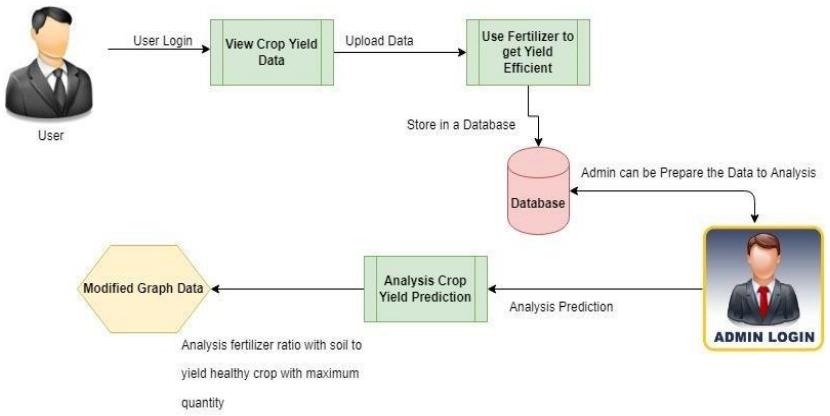


Fig. 5.1: System Design

User (Farmer):

- Farmer (User) need to login into the system using his credentials.
- If the user is new, he/she need to register by feeding their personal details into thesystem.
- After registering he/she can login into the system using their credentials.
- So, once he/she logged into the system, he/she can view crop details based soil typeand area.
- The farmer (user) can also add the crop details if suppose he is having the maximumproduction.
- The fertilizer ratios data will also be feed into the system while the farmer feeding thecrop details.
- All the data feed by the user will be stored in the database.

Admin (Agricultural Officer):

- Admin will also carryon same login procedure as above mentioned in the usersection.
- Admin can analyze the crop yield prediction data by logging into the system.
- Admin can also analyze the fertilizer ratio with soil to yield healthy crop withmaximum quantity.
- Admin can analyze farmer queries using the graphs with respect to districts.

5.2 Data Flow Diagram

5.2.1 Admin Data Flow

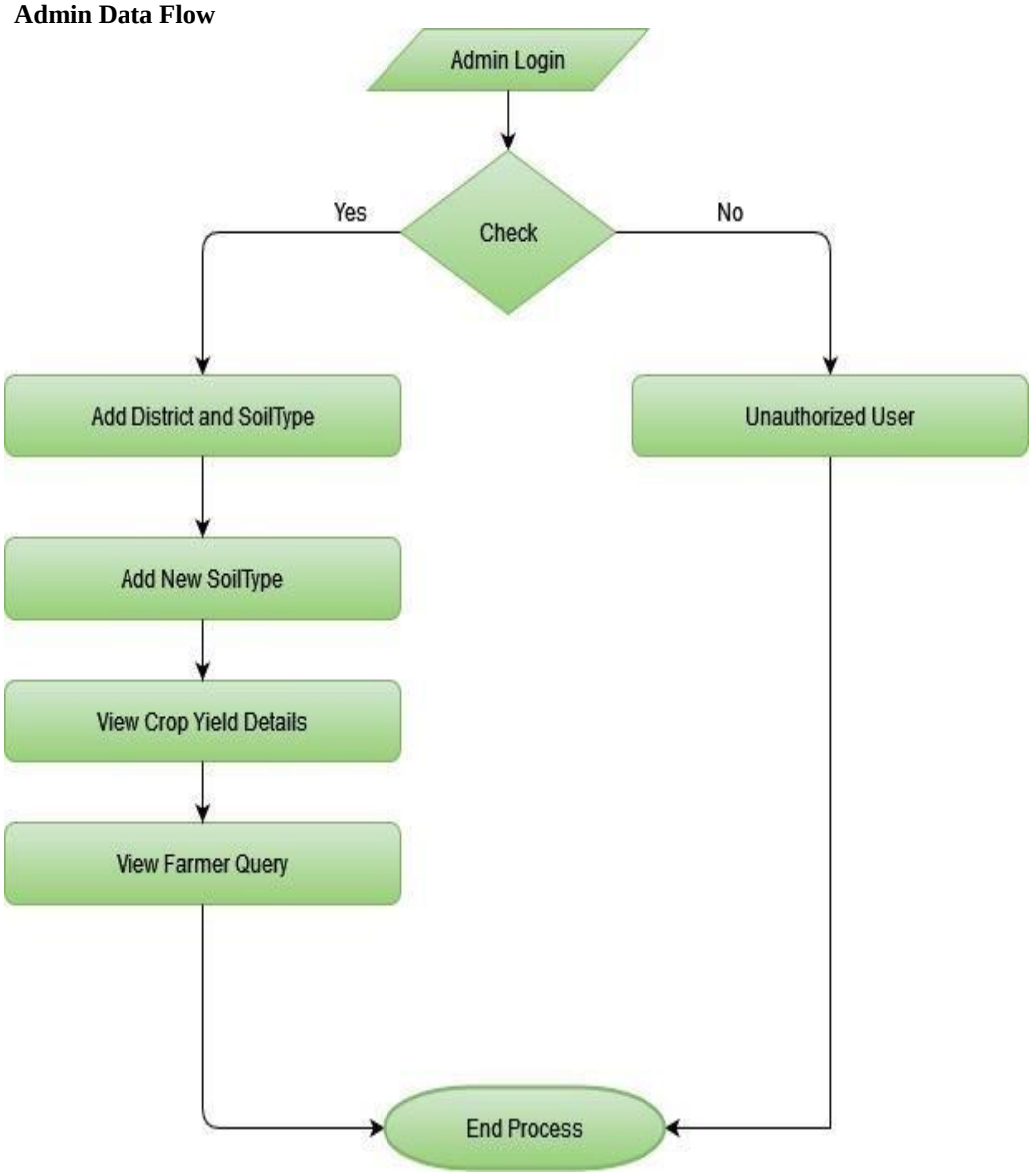


Fig. 5.2: Admin (Agricultural officer)

Steps:

i. The admin should login into the system.
ii. Once logged in, the admin can access the system.
iii. If the admin unmatched the username and password, the system would not allowthe admin.

5.2.2 User Data Flow

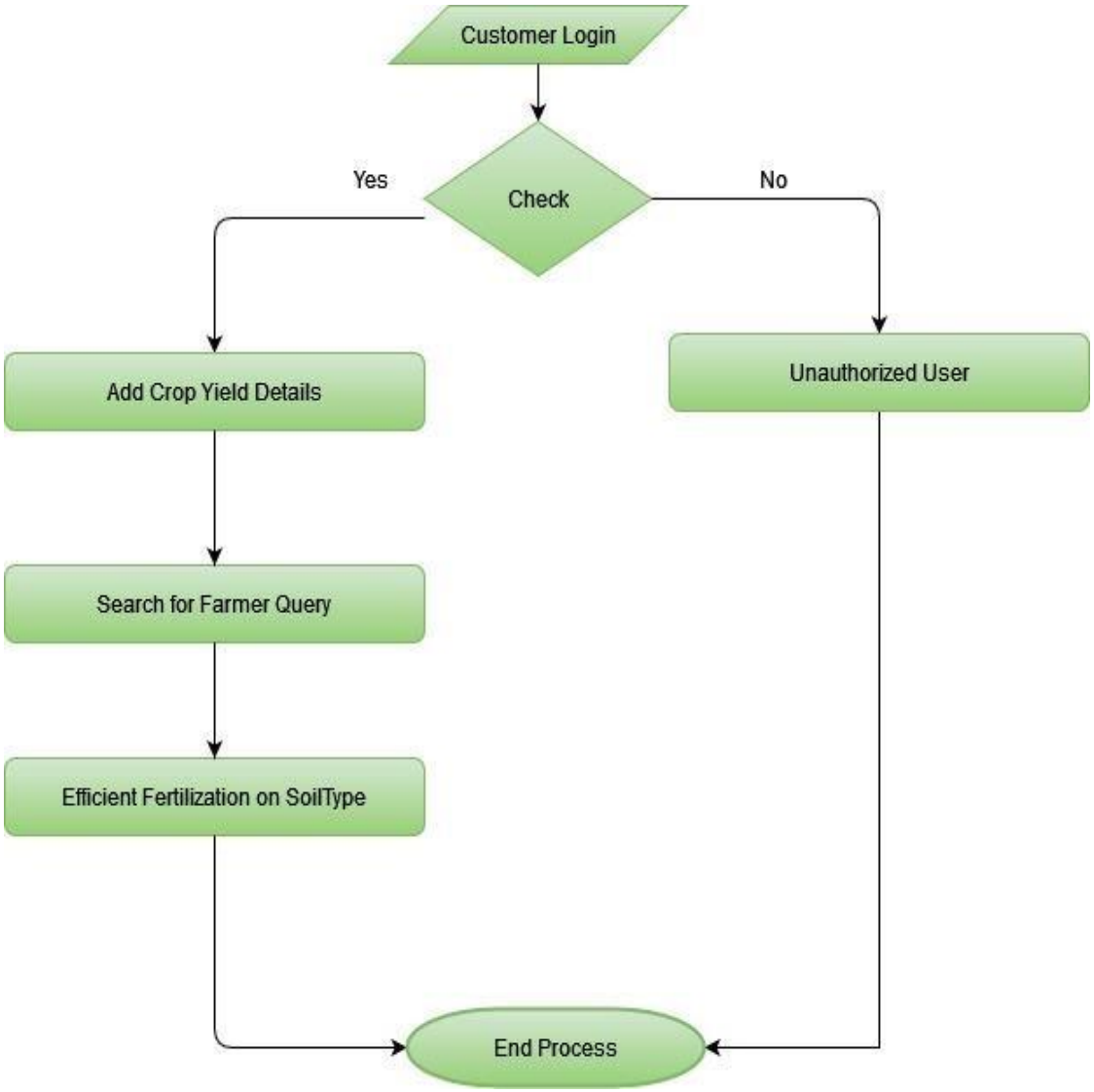


Fig. 5.3: User (farmer)

Steps:

i. The user should login using username and password into the system.
ii. Once the user logged in, user can access the system.
iii. Hence, the user can edit the data into the system and use the system.

5.3 Use Case Diagram

5.3.1 Admin Use Case

The Use case diagram here is the way of summarizing details of the web application whichpredicts the crop and the user (farmer) within the system.

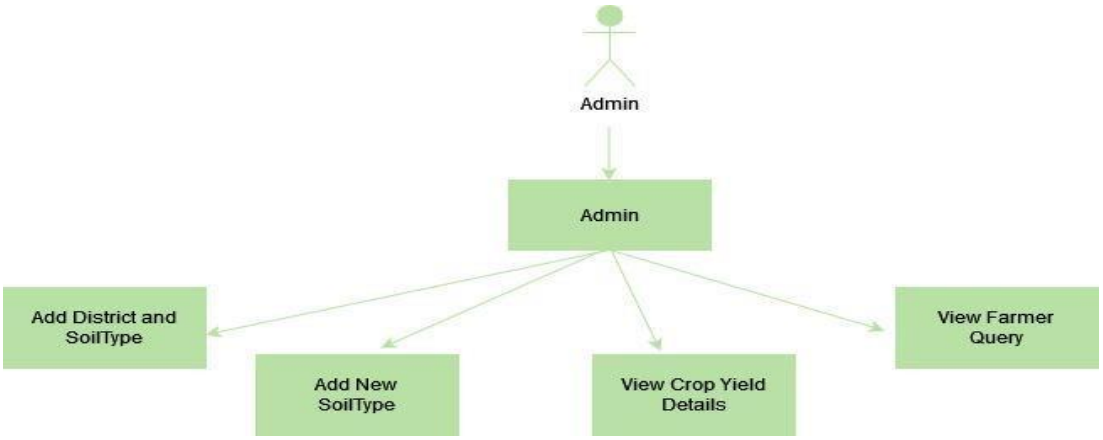


Fig. 5.4: Admin

The fig 5.4 illustrates an Admin use case in which admin can login with his properusername and password. Only admin has right to login system.

5.3.2 User Use Case

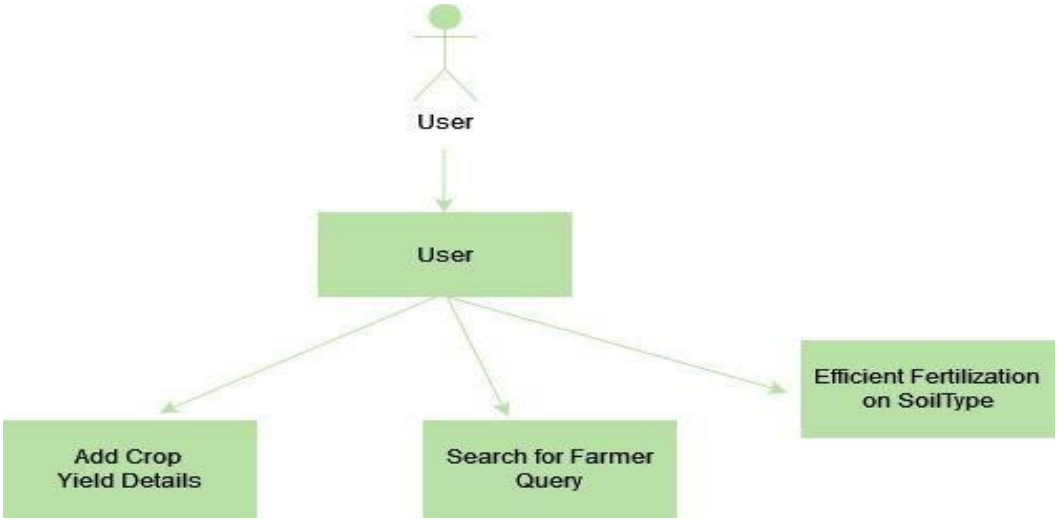


Fig. 5.5: User

The fig 5.5 illustrates the user use case in which first user will register and login with hisproper username and password.

5.4 Activity Diagram

5.4.1 Admin

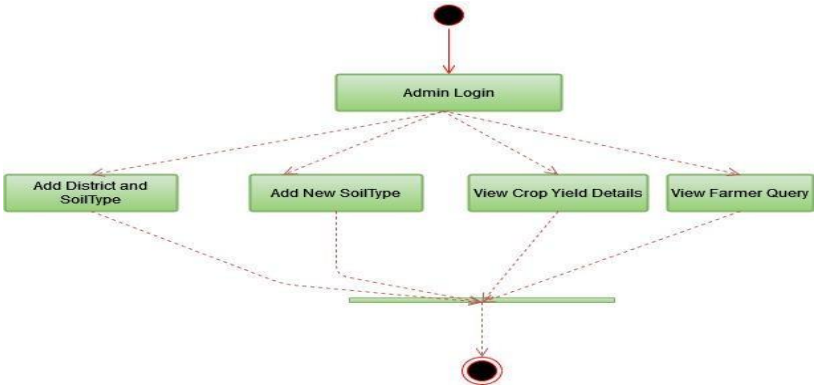


Fig. 5.6: Admin

The fig.5.6 is basically a flowchart to represent the flow from one activity to anotheractivity in case of admin; admin section can perform following activities

- Add district and soil type
- Add new soil type
- view crop yield details
- view farmer query

5.4.2 User

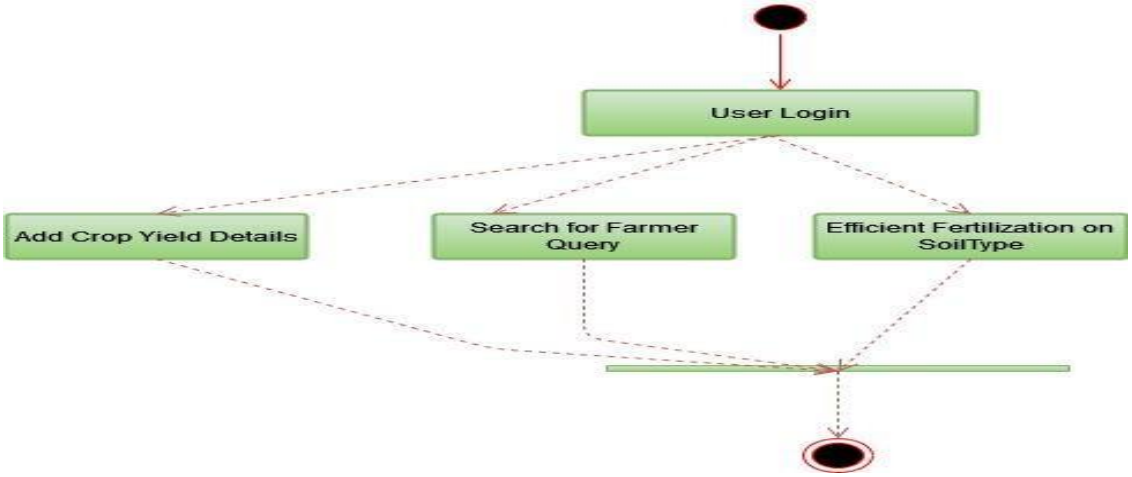


Fig. 5.7: User

The fig.5.7 is basically a flowchart to represent the flow from one activity to anotheractivity in case of user. User can perform following Activities

- Add crop yield details
- Search for farmer query
- Efficient fertilization on soil type

VI. IMPLEMENTATION

6.1 Data Collection

The proposed system follows K-Nearest Neighbor algorithm, the supervised learning algorithm consist of the four levels to calculated and predict the crop for the suitable climate in phenomenon such as,

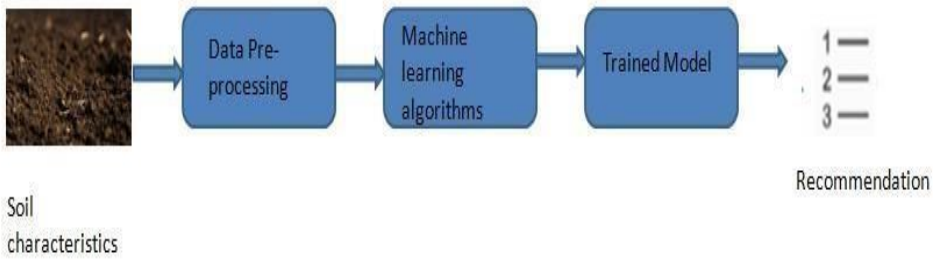


Fig. 6.1: Data flow diagram

Data is composed from a different source and optimized for data sets. And the data is used to evaluate descriptively. Several abstract online outlets, like Kaggle, Google weather forestation and data government, provide the data for up to 10years in series. The data sets such as soil nature, climatic conditions and seed data are used for the crop prediction and better crop yields.

6.2 Preprocessing Step

Preprocessing the data is considered as a significant step machine learning phase. Preprocessing involves adding the missing values, the correct set of data, and extracting the functionality .Dataset form is important to the process of analysis. The data collected in this step will induced in Google Collar platform in the form of pythonprogramming in order to get the desired output.

6.3 Feature Extraction

Extraction of the features would reduce the data size involved to characterize a wide collection of data. The characteristics of soil, crop and weather collected from the pretreatment process establish the final training data collection. This approach selects the features based on the correlation matrix.

6.4 Data Prediction

In advance to this step there need to split the data into train dataset and test dataset. Inthis, we consider each of the characteristics in our training set as a different dimension in some space, and take the value an observation has for this characteristic to be its coordinate in that dimension, so getting a set of points in space. We can then consider the similarity of two points to be the distance between them in this space under some appropriate metric. The way in which the algorithm decides which of the points from the training set are similar enough to be considered when choosing the class to predict for a new observation is to pick the k closest data points to the new observation, and to take the most common class among these. This is why it is called the k Nearest Neighbors algorithm. The implementation of algorithm can be noted as below:

- Load the data
- Initialize K to your chosen number of neighbors
- For each example in the data ϖ calculate the distance between the query example and the current example from the data. ϖ Add the distance and the index of to an ordered collection.
- Sort the ordered collection of distances and indices from smallest to largest (in ascending order) by the distances
- Pick the first K entries from the sorted collection
- Get the labels of the selected K entries
- If regression, return the mean of the K labels
- If classification, return the mode of the K labels

Consider parameters like humidity, rainfall, soil type, area etc. The farmer assigns location, area, soil type as input parameters although other parameters. The crop yield which is an unknown value can be predicted using the values of the nearest known neighbors. This is possible by calculation Euclidian distance between those points. Thus the system would predict crop yield for the given input parameters. The calculation of distance between points in a feature space, different distance functions could be used, in which the Euclidean distance function is the most commonly used one. Say p and q are represented as feature vectors. To measure the distance between p and q, the Euclidean metric is generally used by if a = (a1, a2) and b = (b1, b2) then the distance is given by:

d (a, b) = $\sqrt{(b1 - a1)^2 + (b2 - a2)^2}$.

VII. TEST CASES: VALIDATING AND VERIFYING THESOFTWARE

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

7.1 Types of testing

7.1.1 Unit testing:

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

7.1.2 Integration testing:

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

7.1.3 Functional testing:

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals. Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.
Invalid Input : identified classes of invalid input must be rejected.
Functions : identified functions must be exercised.
Output : identified classes of application outputs must be exercised.Systems/Procedures : interfacing systems or procedures must be invoked.

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

7.1.4 System testing:

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

7.2 Test Cases:

Test cases include input data, execution condition, and prediction criteria. It consists of theprogram input and expected output.

7.2.1 Registration Form:

Table 7.1: Registration Form

SI.NO.	Input Variable	Expected Result	Actual Result
1.	First Name only andRegister.	Error Message. (All the required fields need to be filled.)	Display an errormessage.
2.	Last Name only and Register.	Error Message. (All the required fields need to be filled.)	Display an errormessage.
3.	First Name and Last Nameonly and Register.	Error Message. (All the required fields need to be filled.)	Display an errormessage.
4.	Contact Number less than 10digits and Register.	Error Message. (Contact Number must be of 10- digit number and all the required fields need to be filled.)	Display an errormessage.
5.	Password only and Register.	Error Message. (All the required fields need to be filled.)	Display an errormessage.
6.	First name, Last name, Password and 10-digit contact number.	Candidate should beable to Register.	Candidate Register successfully.

7.2.2 : Login Form:

Table 7.2: Login Form

SI.NO.	Input Variable	Expected Result	Actual Result
1.	Username only andLogin.	Error Massage (both Usernameand password must be entered)	Display an error message.
2.	Password only andLogin	Error Massage (both Usernameand password must be entered)	Display an error message.
3.	Username andPassword and Login.	If valid then continue else anerror message for invalid user.	User Successfullylogged in.

7.2.3 : Unknown Inputs:

Table 7.3: Unknown Inputs

SI.NO.	Input Variable	Expected Result	Actual Result
1.	Unknown Soil type	Prediction results (Should enterknown soil type).	No results arepredicted.
2.	Unknown district	Prediction results (Should enterknown district)	No results arepredicted.

VIII. RESULT AND ANALYSIS

In this chapter system functionality modules are showcased using snapshots.

8.1 Home page:

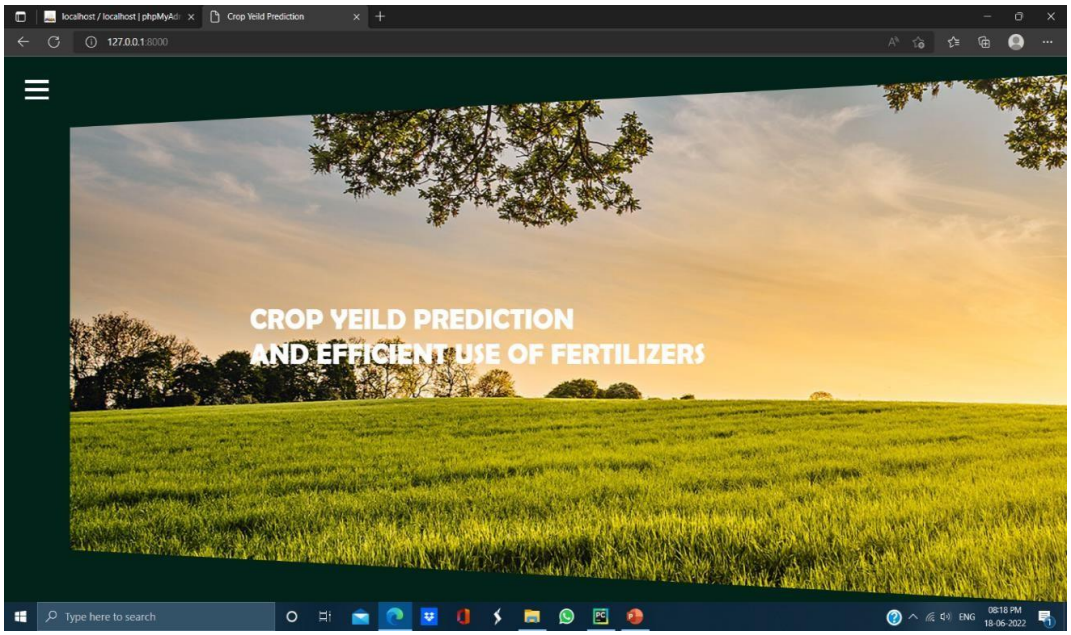


Fig. 8.1: Home page

This is the first and foremost page of the system which consists the title of the web application. Here we can find the options at the left corner of the web page to proceed the process.

8.2 Window in homepage to choose:

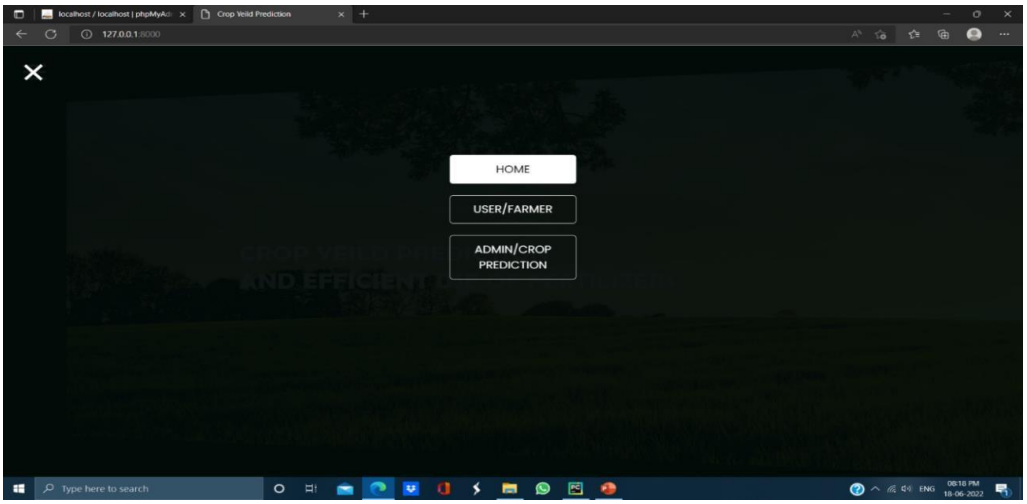


Fig. 8.2: Window in homepage to choose

The fig 8.2 shows the system in which a window is opened by clicking into the left corner of the homepage, show cases the options to choose whomever want to access the system. Here we find options like User and Admin.

8.3 Login Page:

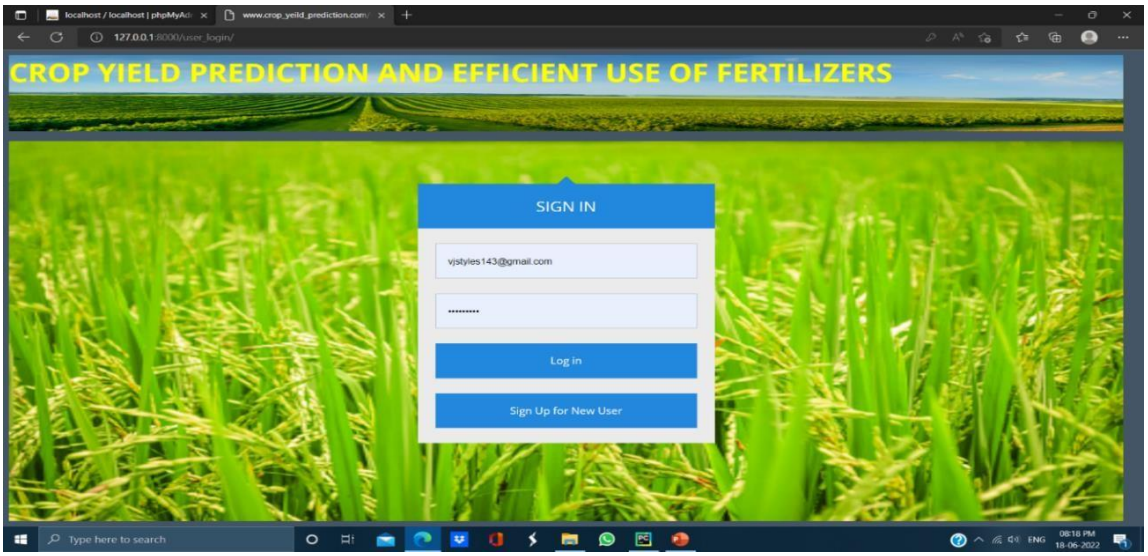


Fig. 8.3: Login Page

The fig 8.3 which appears on clicking the user button in the precious page. This is the login page to access the system for either user or admin using their credentials like username etc.

8.4 Sign up page:

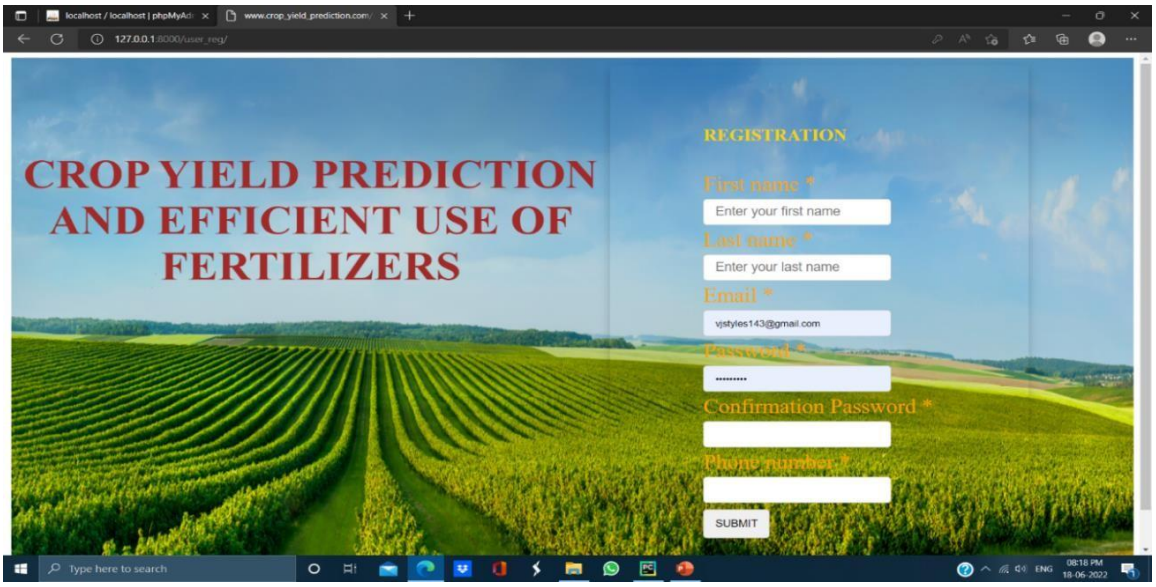


Fig. 8.4: Sign up page

The fig 8.4 occurs when farmer sign up with their details into the system. The signup page for farmer where farmer needs to enter all farmer personal credentials like first name, last name, email address password mobile number etc.

8.5 Window for user/farmer to choose:

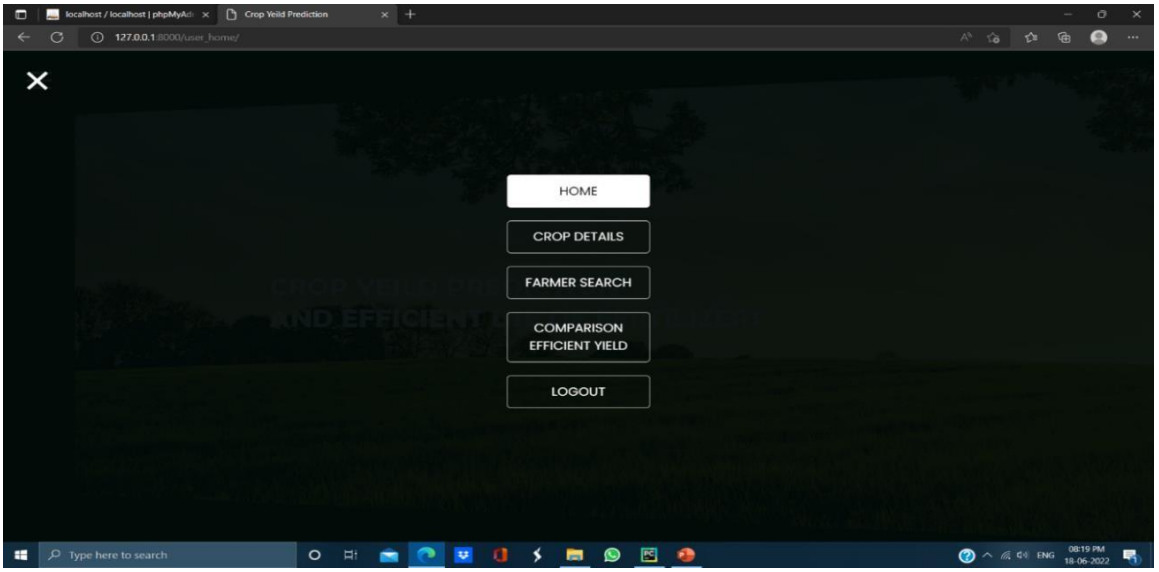


Fig. 8.5: Window for user/farmer to choose

The fig 8.5 is another sub main page which occurs after clicking on the admin. Anotherwindow to choose sections in the user part where multiple options showcased.

8.6 Crop details page:

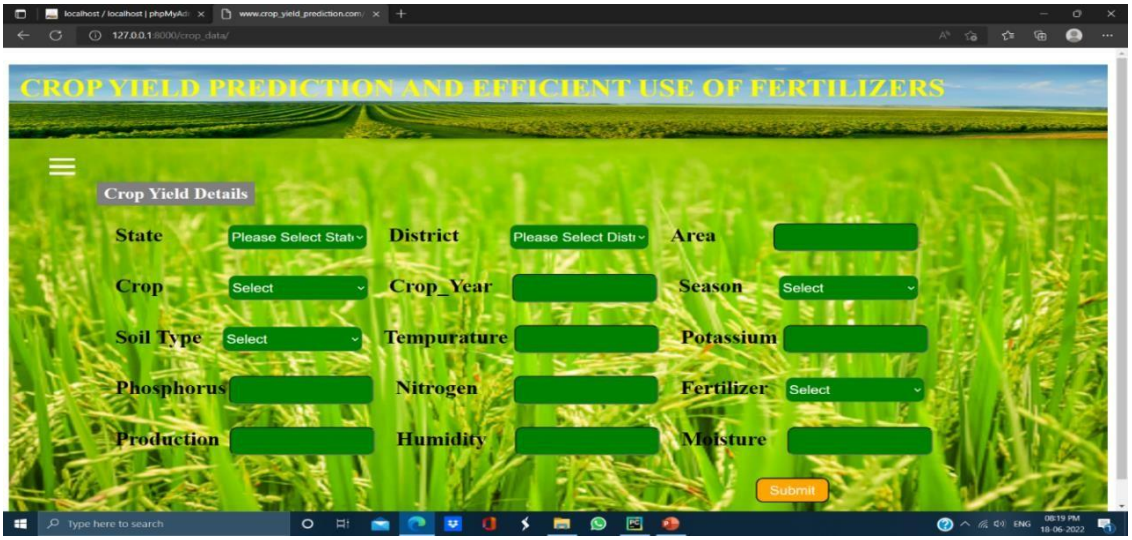


Fig. 8.6: Crop details page

The fig 8.6 shows crop details page where user need to enter all the details of the crop farmer chooses to seed like state, district, area of land, crop, crop year, season, soiltype, temperature, fertilizer etc.

8.7 Farmer search page:

Fig. 8.7: Farmer search page

The fig 8.7 is farmer search page where the user searches crop using the kind of the soil he has in the land whether the soil is of either clay or black or something else kind of soil.

8.8 Comparison yield efficient page:

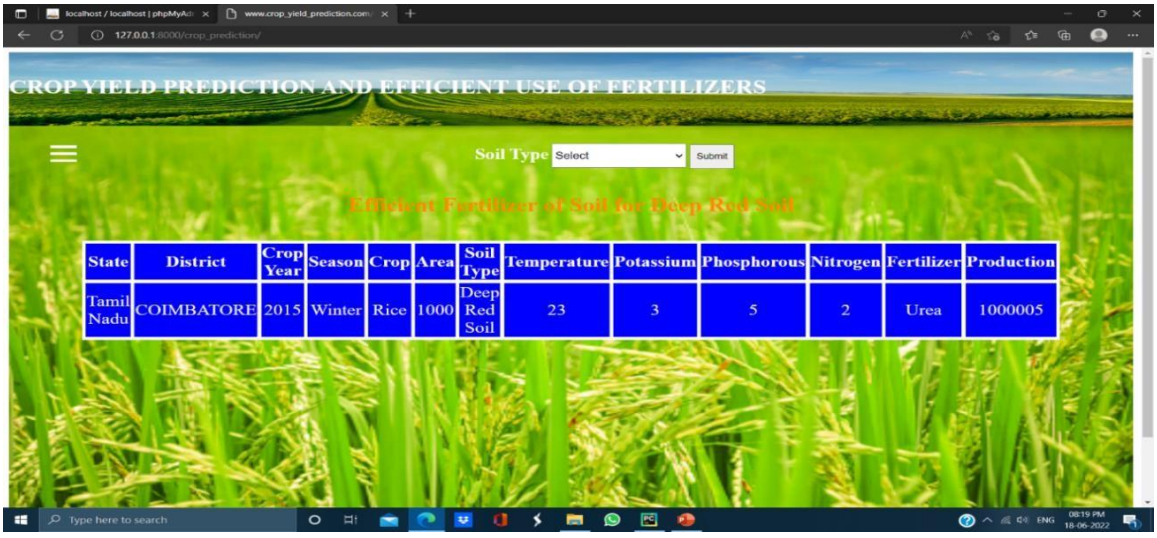


Fig. 8.8: Comparison yield efficient page

The fig 8.8 is the efficient fertilizer page and this page gives the efficient fertilizer for whatever the type of soil user has entered, fertilizers like DAP, Urea, etc.

8.9 Signup page for admin:

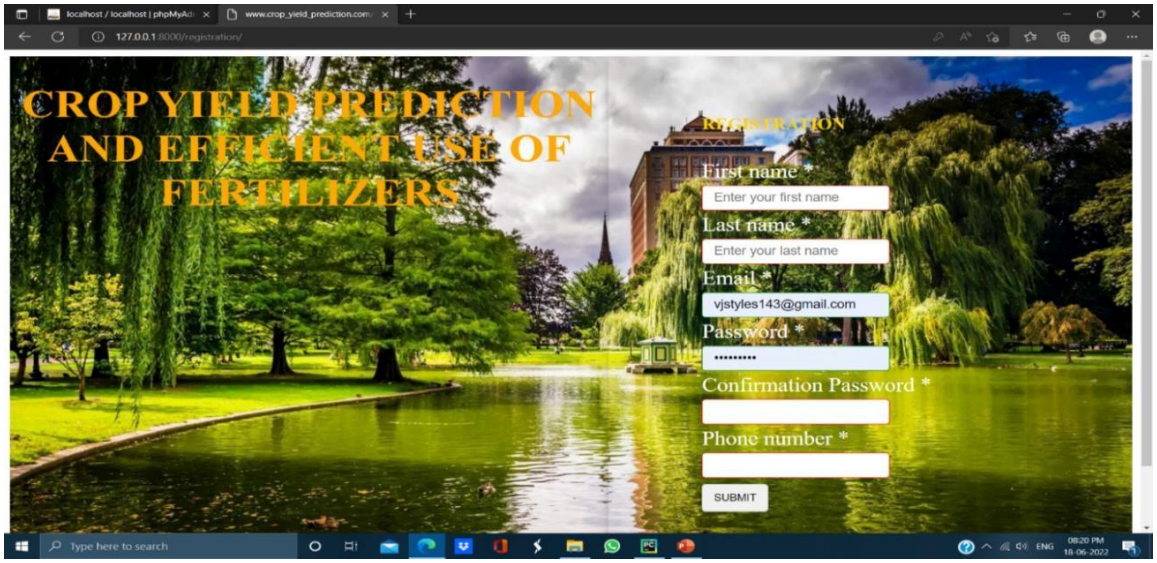


Fig. 8.9: Signup page for admin

The fig 8.9 is the sign-up page occurs after clicking on the admin sign up button. The signup page for admin to enter the system by logging into the system by filling details like first name, last name, email address, password, phone number etc.

8.10 Window for admin to choose:

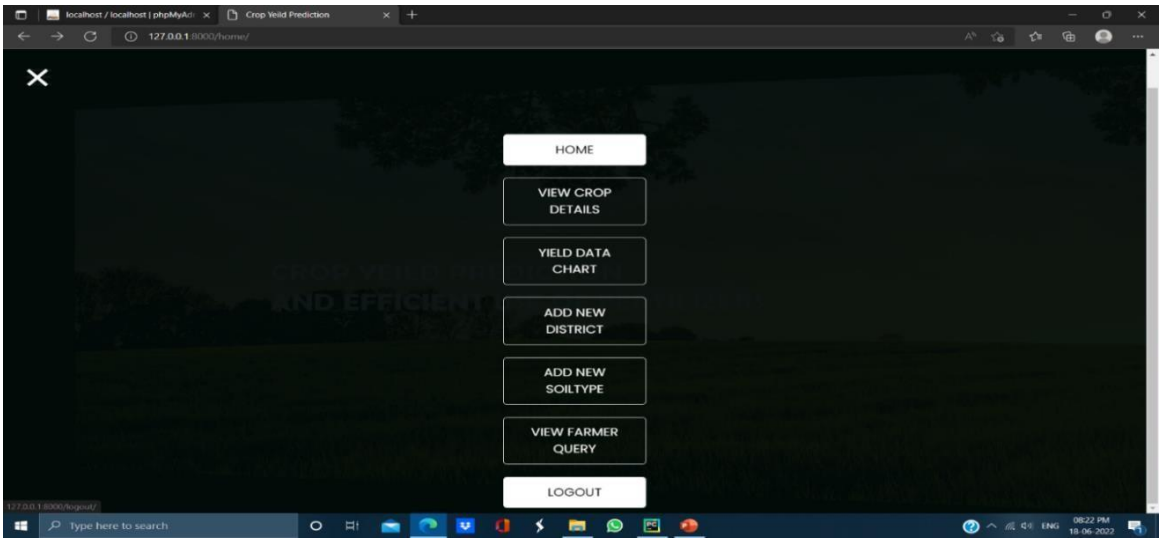
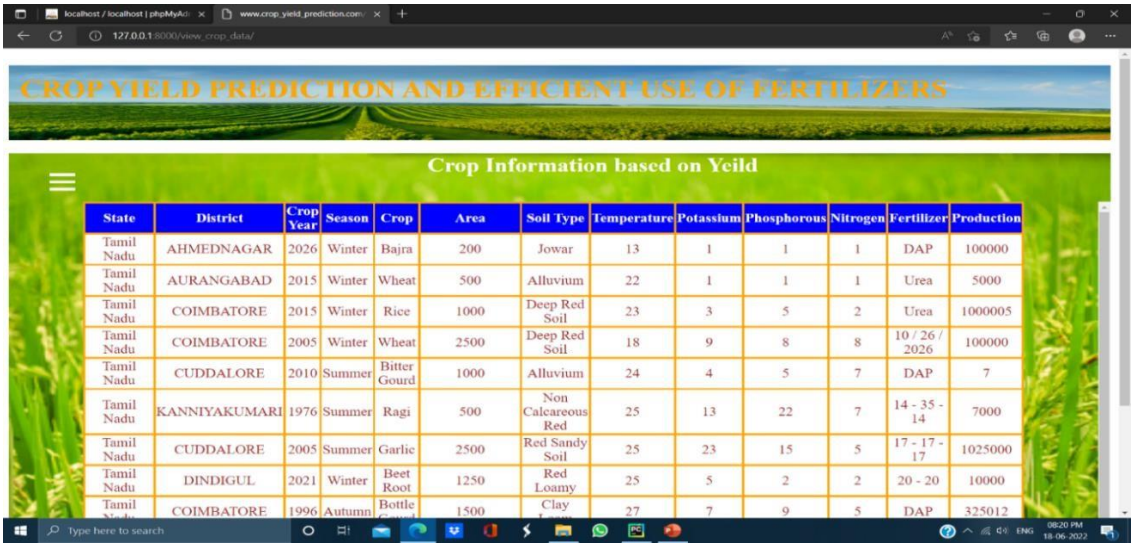


Fig. 8.10: Window for admin to choose

The fig 8.10 is the page occurs on clicking the admin option at the main page of the website. The window to choose the different options for admin comes after clicking admin button.

8.11 Crop details page:



State	District	Crop Year	Season	Crop	Area	Soil Type	Temperature	Potassium	Phosphorous	Nitrogen	Fertilizer	Production
Tamil Nadu	AHMEDNAGAR	2026	Winter	Bajra	200	Jowar	13	1	1	1	DAP	100000
Tamil Nadu	AURANGABAD	2015	Winter	Wheat	500	Alluvium	22	1	1	1	Urea	5000
Tamil Nadu	COIMBATORE	2015	Winter	Rice	1000	Deep Red Soil	23	3	5	2	Urea	1000005
Tamil Nadu	COIMBATORE	2005	Winter	Wheat	2500	Deep Red Soil	18	9	8	8	10 / 26 / 2026	100000
Tamil Nadu	CUDDALORE	2010	Summer	Bitter Gourd	1000	Alluvium	24	4	5	7	DAP	7
Tamil Nadu	KANNIYAKUMARI	1976	Summer	Ragi	500	Non Calcareous Red	25	13	22	7	14 - 35 - 14	7000
Tamil Nadu	CUDDALORE	2005	Summer	Garlic	2500	Red Sandy Soil	25	23	15	5	17 - 17 - 17	1025000
Tamil Nadu	DINDIGUL	2021	Winter	Beet Root	1250	Red Loamy	25	5	2	2	20 - 20	10000
Tamil Nadu	COIMBATORE	1996	Autumn	Bottle Gourd	1500	Clay Loamy	27	7	9	5	DAP	325012

Fig. 8.11: Crop details page

The fig 8.11 occurs when admin choose to view the crop information. The crop details page for admin to analyze the prediction and the details here come from the user entered data of the crop.

8.12 Data chart page:

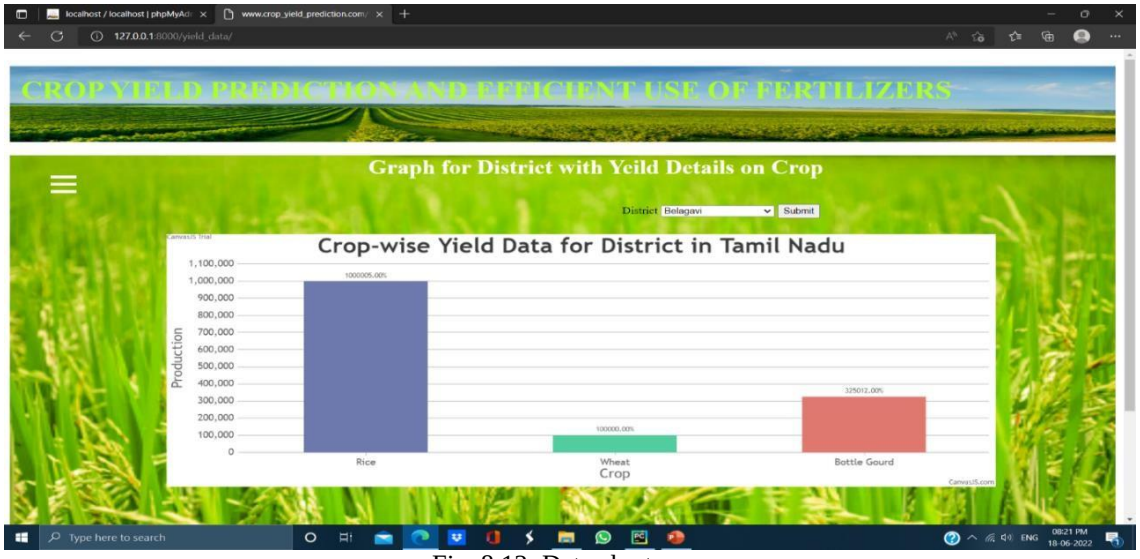


Fig. 8.12: Data chart page

The fig 8.12 is the second optional page of admin section. The data chart where itdepicts the crop-wise yield data for particular district the user opts for in the system.

8.13 Add new district page:

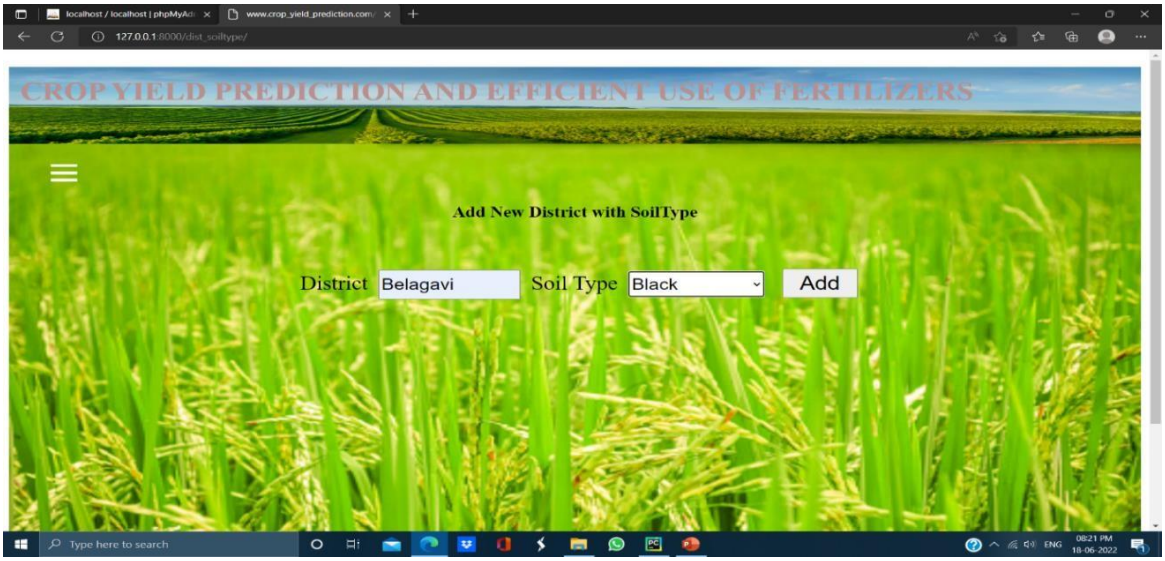


Fig. 8.13: Add new district page

The fig 8.13 is data additional page. The page here is to add a new district into the systemby admin if needed to enter occasionally. Whenever a new district is been made into the state then this option comes to use.

8.14 Add new soil type page:

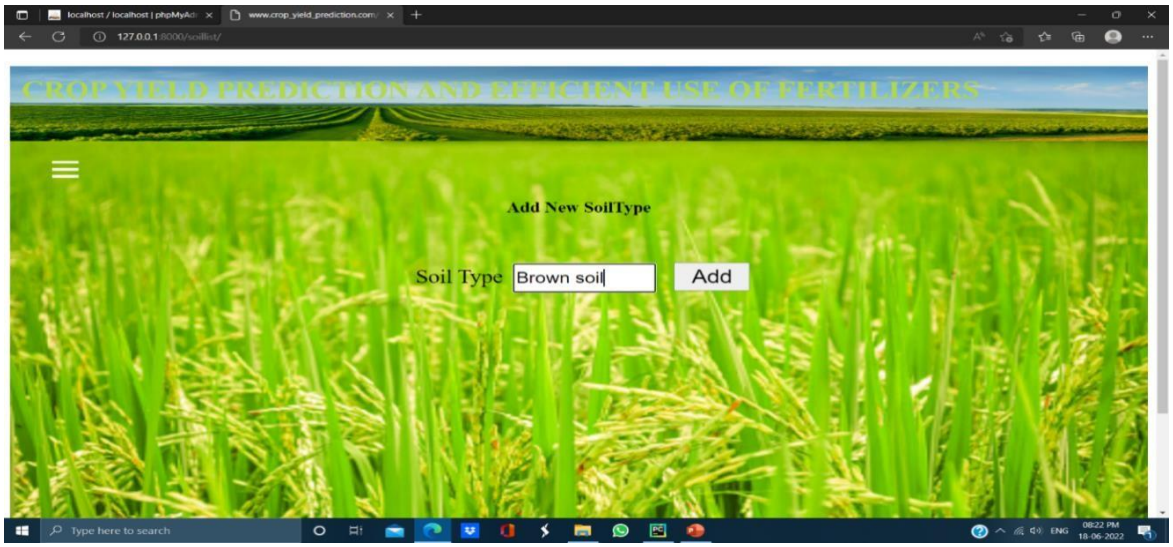


Fig. 8.14: Add new soil type page

The fig 8.14 is also a data addition page. The page here is to enter new soil type if the admin found to add the new kind soil into the system whether informed or self- researched.

8.15 Farmer query page:

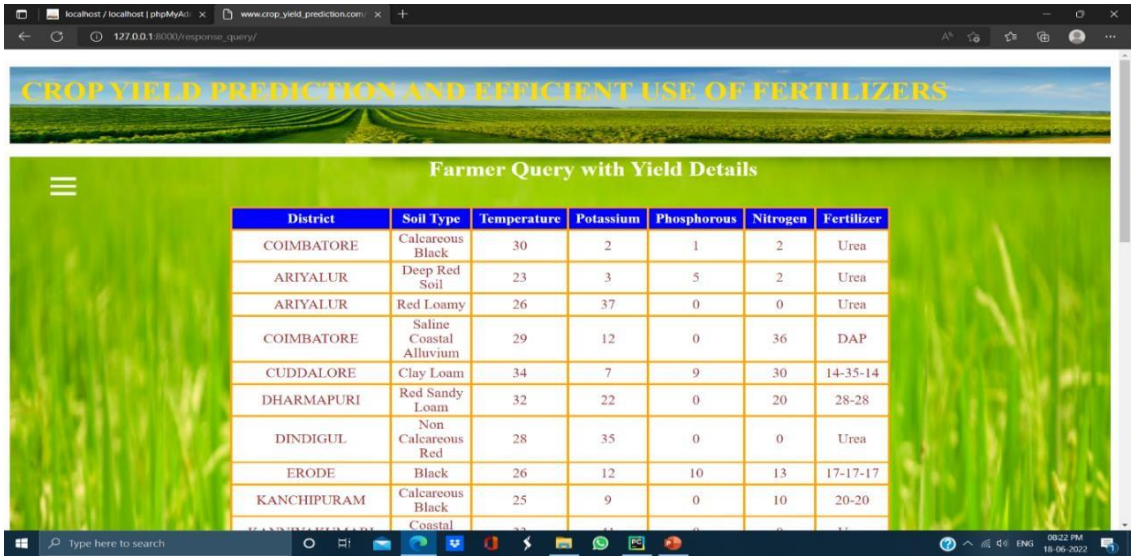


Fig. 8.15: Farmer query page

The fig 8.15 is the farmer query details viewing page. The farmer query page willshowcase all the queries enter by the user into the system by logging into the system.

8.16 Final prediction of crop:

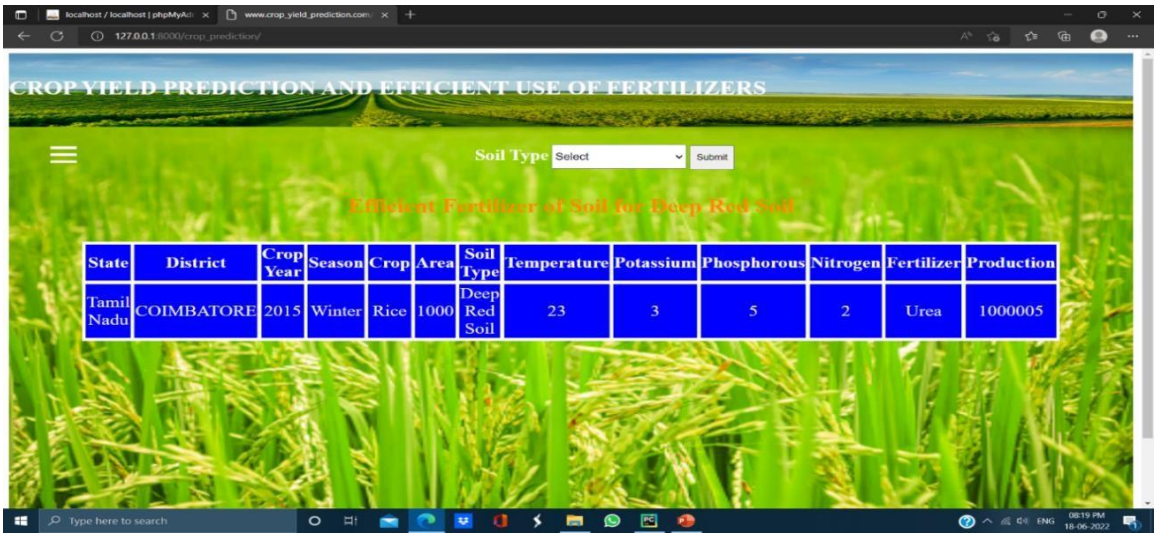


Fig. 8.16: Final prediction of crop

The fig 8.16 is the final prediction of crop where farmer gets all the required predictedvalues like fertilizer ratio and production cost with farmer belonging state.

IX. CONCLUSION AND FUTURE SCOPE

Conclusion

The proposed system predicts the appropriate crop for the farmer taking in considerations of all the factors which includes soil type, pH value, humidity, rainfall, and fertilizer ratio. The system experimented with different types of crops in various regions across India to predict the output. Even fertilizer data are trained using the KNN algorithm and evaluated to get the proper result by using nitrogen, phosphorus etc. The model for the crop production is compared in predicting the output and by various parameters with respect to the error rate. While predicting the output for the model and the compassion is plotted in the graph.

Future scope

- In coming years, can try applying data independent system. Integrating soil details to the system is an advantage, as for the selection of crops knowledge on soil is also a parameter.
- Proper irrigation is also needed feature crop cultivation. In reference to rainfall can depict whether extra water availability is needed or not. This research work can be enhanced to higher level by availing it to whole India.

Python:

APPENDIX A: TOOLS USED

Python is an interpreter, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can befreely distributed.

Often, programmers fall in love with Python because of the increased productivity it provides. Since there is no compilation step, the edit-test-debug cycle is incredibly fast. Debugging Python programs is easy: a bug or bad input will never cause a segmentation fault. Instead, when the interpreter discovers an error, it raises an exception. When the program doesn't catch the exception, the interpreter prints a stack trace. A source level debugger allows inspection of local and global variables, evaluation of arbitrary expressions, setting breakpoints, stepping through the code a line at a time, and so on. The debugger is written in Python itself, testifying to Python's introspective power. On the other hand, often the quickest way to debug a program is to add a few print statements to the source: the fast edit-test-debug cycle makes this simple approach very effective.

Django:

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent the wheel. It's free and open source. Django's primary goal is to ease the creationof complex, database-driven websites. Django emphasizes reusability and "pluggability" of components, rapid development, and the principle of don't repeat yourself. Python is used throughout, even for settings files and data models. If we compare Django with other open-source technologies, it offers the best documentation in the market. Better documentation of any technology is like a very well-established library for any developer.

There, he can search for any function desired with ease with the time involving in the searching purpose only. The documentation of any technology is also one of the categoriesto grade a technology, as it lets other developers other than its own creators to efficiently utilize the technology. Django has been best at documentation from the beginning, from the point it became open source in 2005 to the present date, and the documentation has only been getting better with active development of technology and it is also offered in different languages.Django also provides an optional administrative create, read, update and delete interface that is generated dynamically through introspection and configuredvia admin models.

MySQL:

MySQL is a relational database management system based on the Structured Query Language, which is the popular language for accessing and managing the records in the database. MySQL is open-source and free software under the GNU license. It is supported by Oracle Company. Our MySQL tutorial includes all topics of MySQL database that provides for how to manage database and to manipulate data with the help of various SQL queries. These queries are: insert records, update records, delete records, select records, create tables, drop tables, etc. There are also given MySQL interview questions to help you better understand the MySQL database. MySQL is currently the most popular database management system software used for managing the relational database. It is open-source database software, which is supported by Oracle Company. It is fast, scalable, and easy to use database management system in comparison with Microsoft SQL Server and Oracle Database. It is commonly used in conjunction with PHP scripts for creating powerful and dynamic server-side or web-based enterprise applications.

MySQL follows the working of Client-Server Architecture. This model is designed for the end-users called clients to access the resources from a central computer known as a server using network services. Here, the clients make requests through a graphical user interface (GUI), and the server will give the desired output as soon as the instructions are matched. The process of MySQL environment is the same as the client-server model.
The core of the MySQL database is the MySQL Server. This server is available as a separate program and responsible for handling all the database instructions, statements, or commands. The working of MySQL database with MySQL Server are as follows:

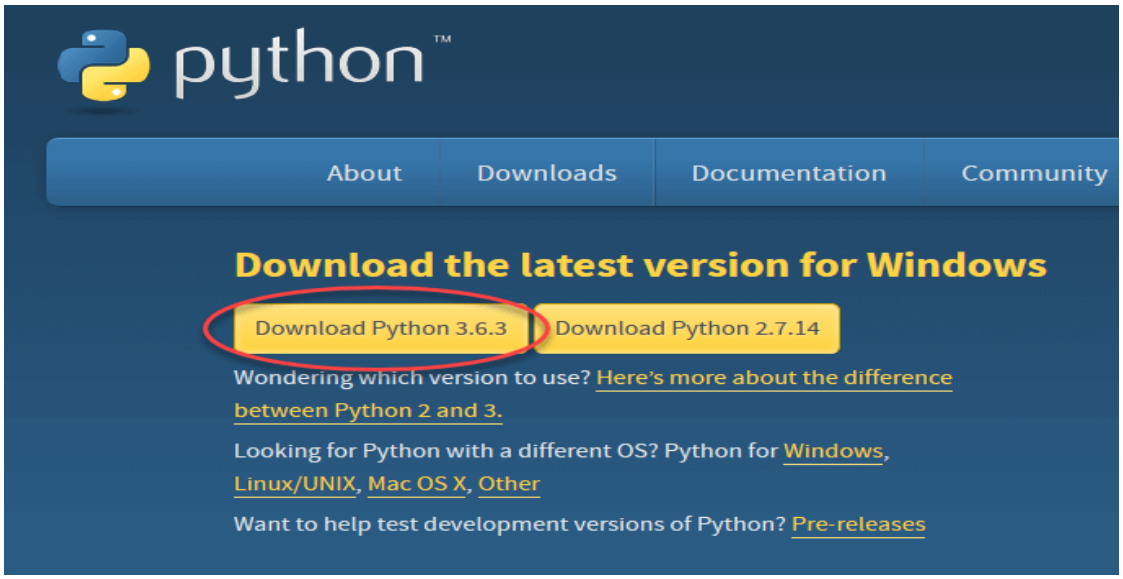
1. MySQL creates a database that allows you to build many tables to store andmanipulate.
2. Clients make requests through the GUI screen or command prompt by using specificSQL expressions on MySQL.
3. Finally, the server application will respond with the requested expressions andproduce the desired result on the client-side.

A client can use any MySQL GUI. But it is making sure that your GUI should be lighter and user-friendly to make your data management activities faster and easier. Some of the most widely used MySQL GUIs are MySQL Workbench, Sequel Ro, DB Visualizer, and the Navicert DB Admin Tool. Some GUIs are commercial, while some are free with limited functionality, and some are only compatible with MacOS. Thus, you can choose the GUI according to your needs.

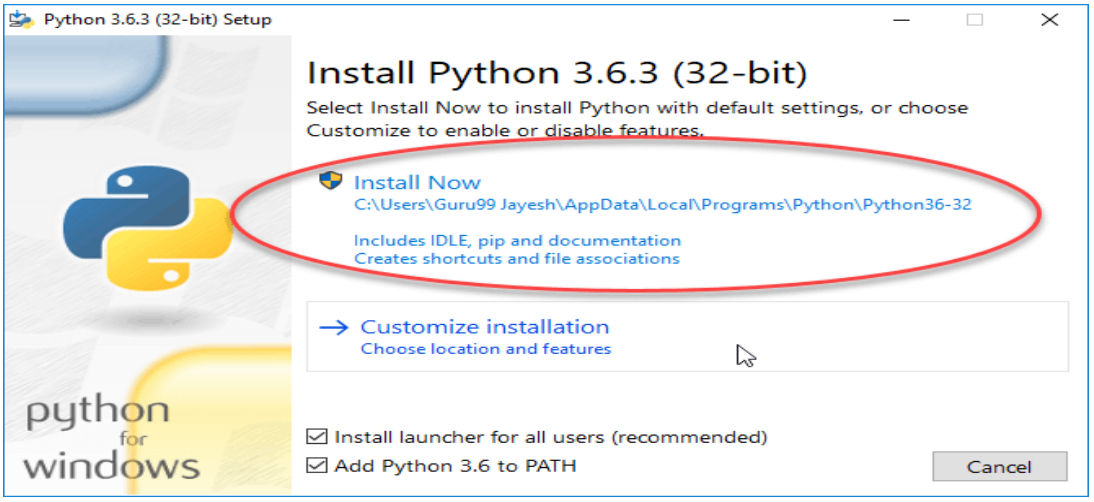
APPENDIX B: MANUAL FOR SOFTWARE

PyCharm installation guide:
PyCharm is a cross-platform editor developed by JetBrains. PyCharm provides all thetools you need for productive Python development.

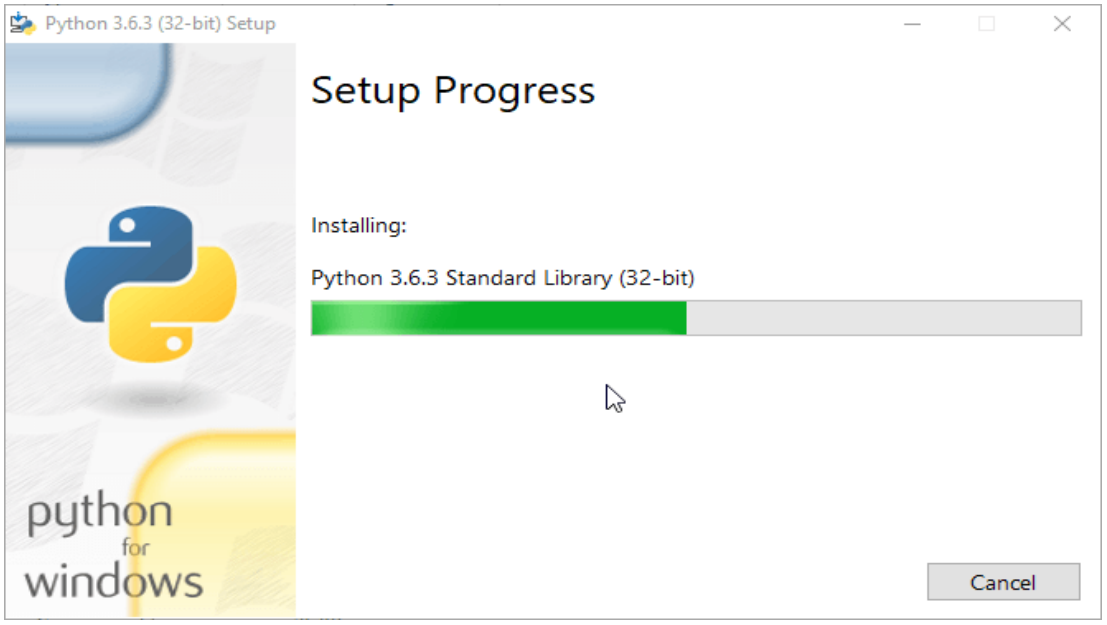
How to install Python IDE:
Step 1: To download python visit website https: //www.python.org/downloads/ andchoose your version. We have chosen Python version 3.6.3.



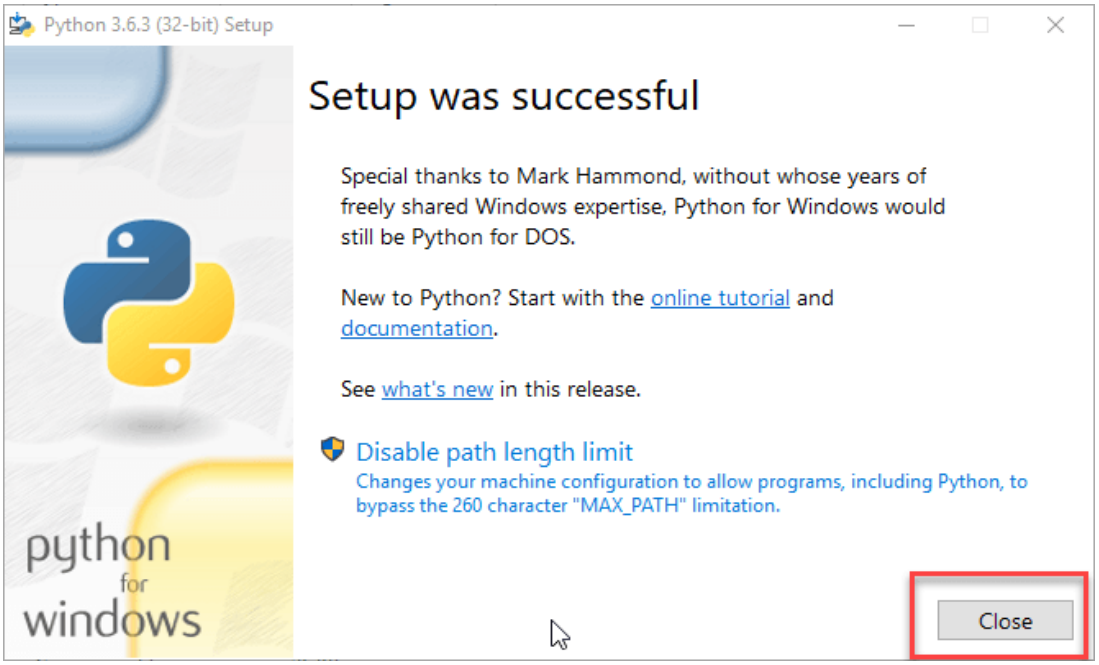
Step 2: Once download file completed run the execute file and open install python.



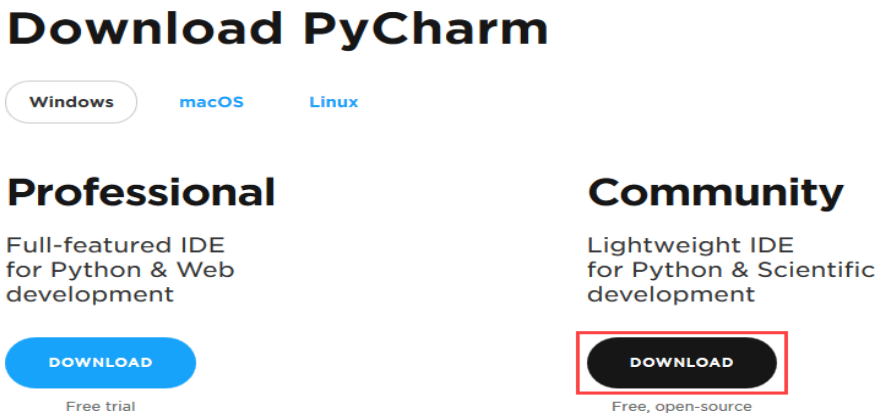
Step 3: Installation starts



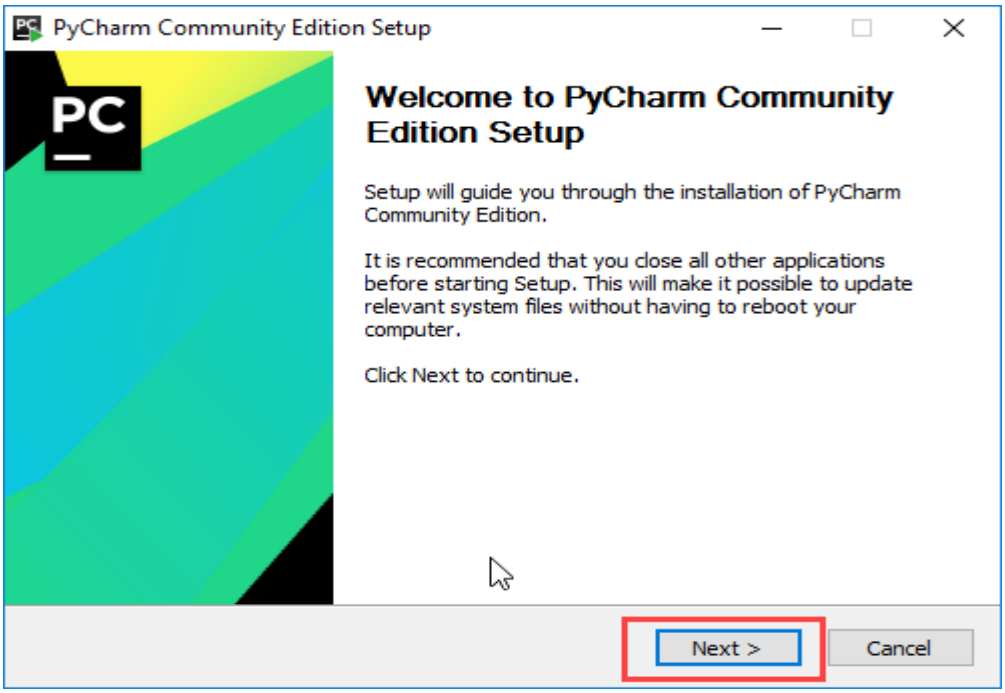
Step 4: when the process completes it shows completed successfully.



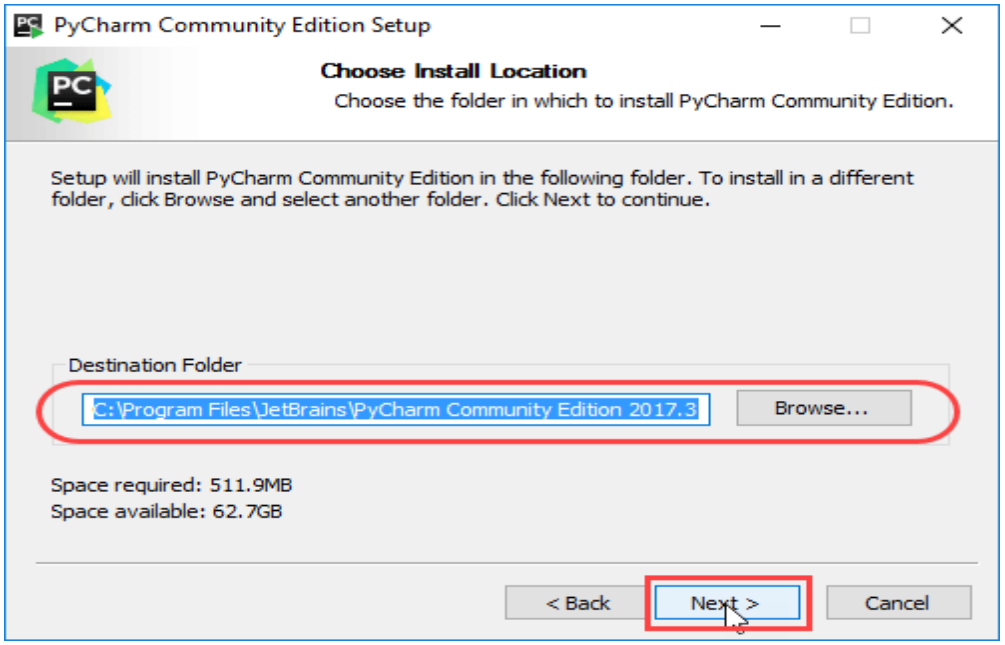
How to install PyCharm:
Step 1: To download website https://www.jetbrains.com/pycharm/download/ and clickthe "download" link under the community sections.



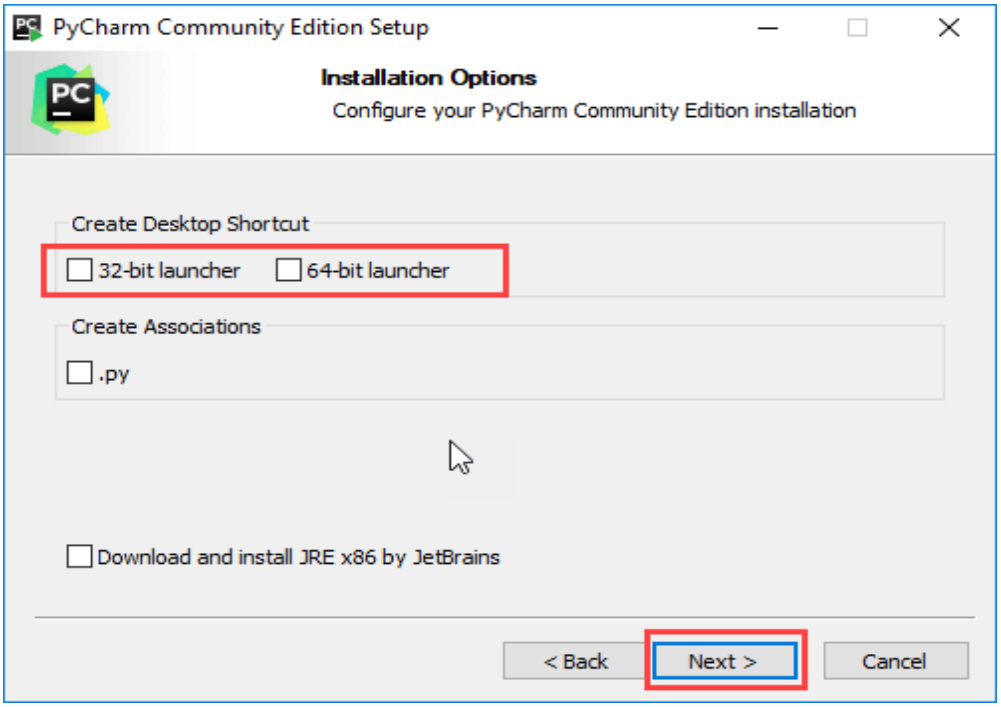
Step 2: Once the download is completed, run the exe for install PyCharm. The setupwizard should have started, click "next".



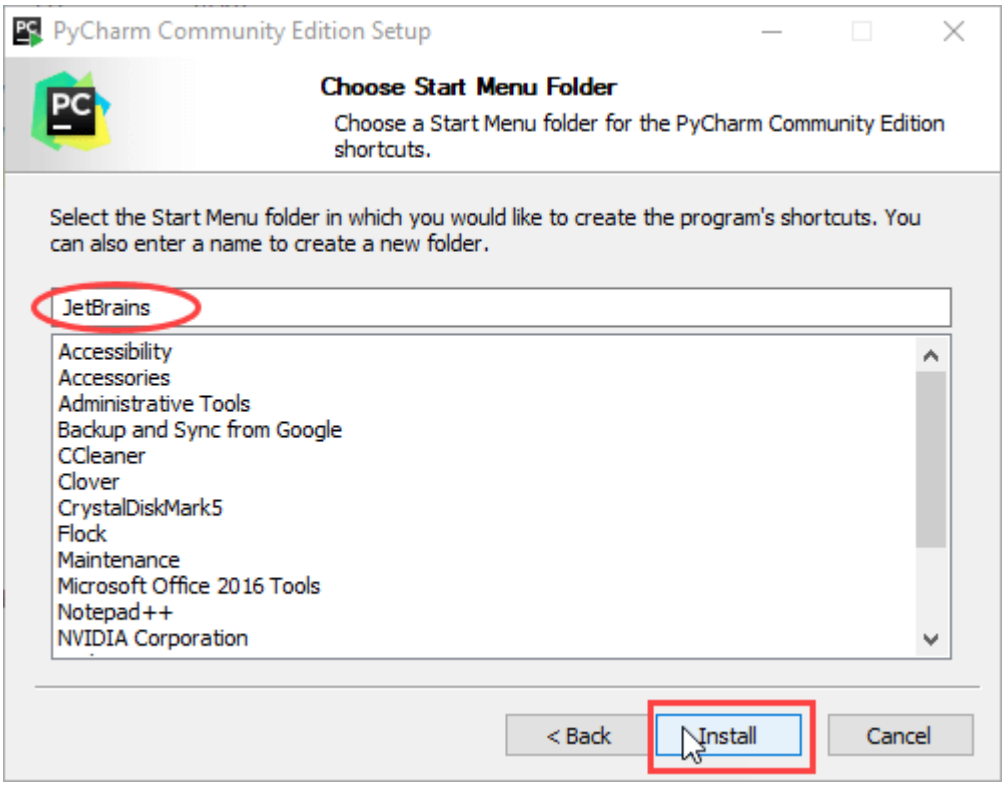
Step 3: On next screen, change the installation path if required. Click "next".



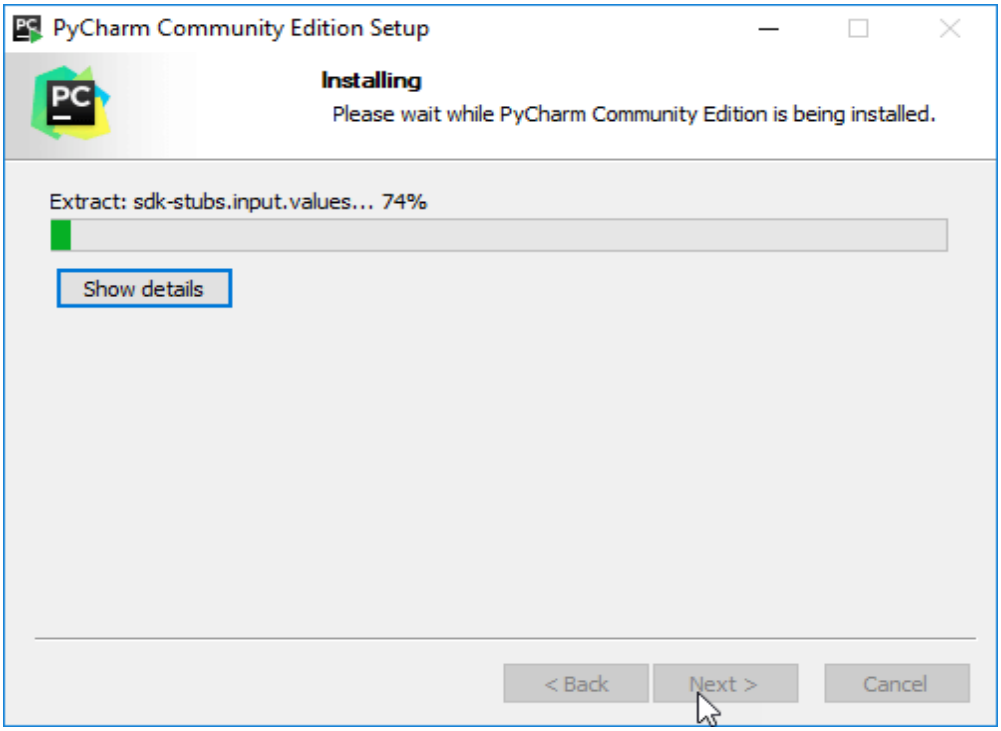
Step 4: On the next screen, you can create desktop shortcut if you want and click onnext.



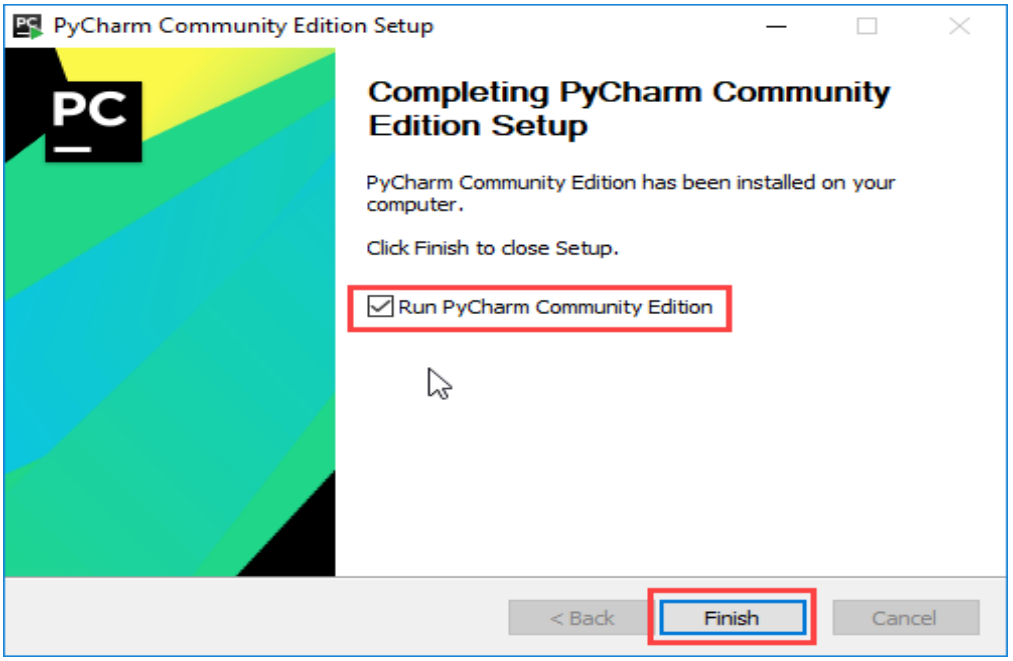
Step 5: Choose the start menu folder and keep selected JetBrains and click on install.



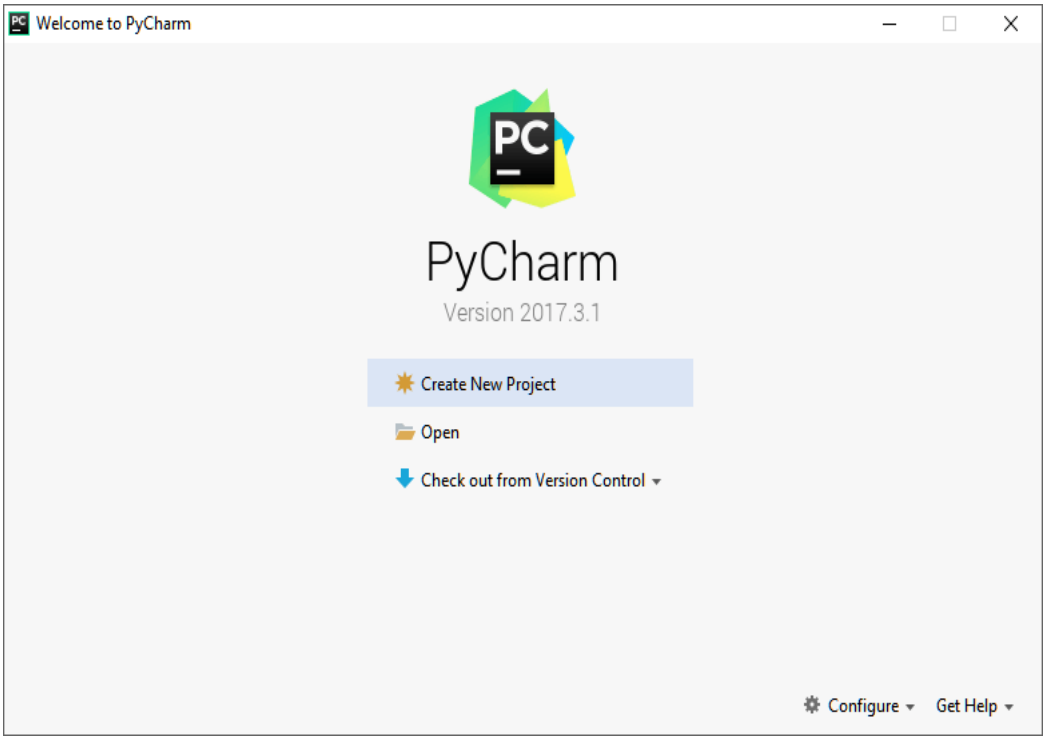
Step 6: Wait for the installation to finish.



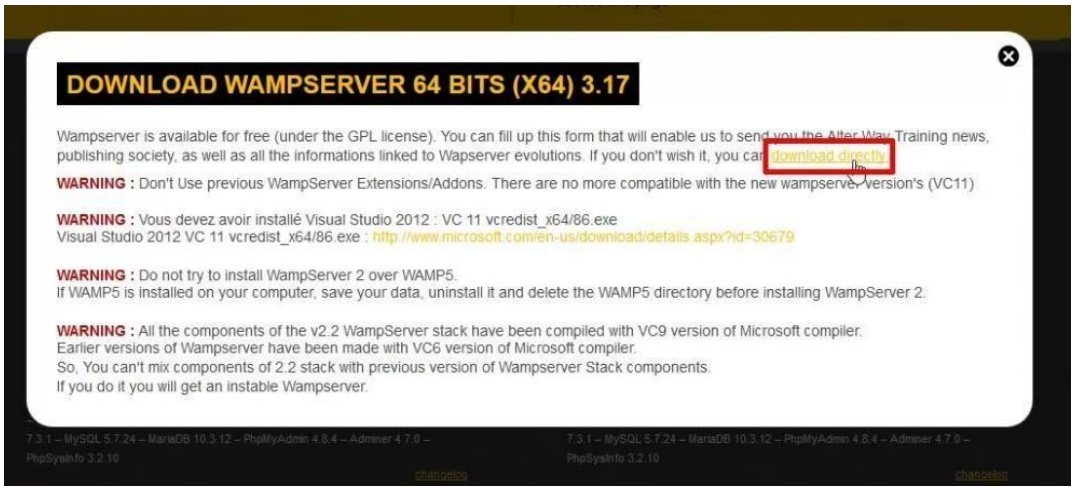
Step 7: Once installation finished, you should receive message screen that PyCharm is installed and if you want to go ahead and run it, click the Run PyCharm Community Edition box first and click on finish.



Step 8: After you click on finish the Following screen will appear.

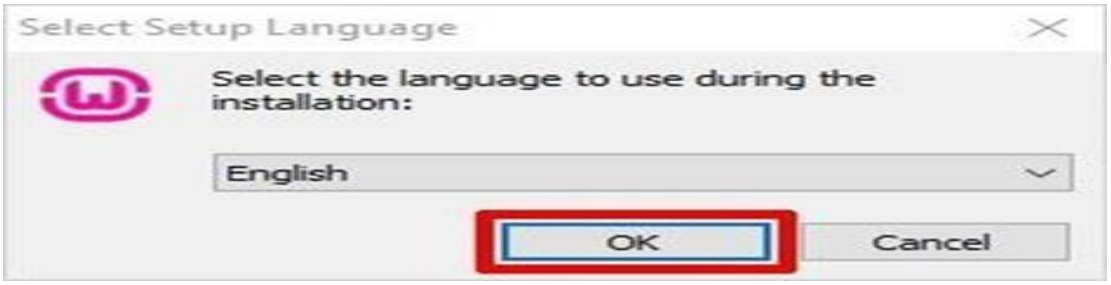


WAMPP installation guide:
Below are the detailed steps on how to install WAMP that you need to follow for asuccessful WAMP installation.
Step 1: Download the WAMP Server
Go to official website of WAMP Server and choose the version according to yourcomputer.



Step 2: Initiate WAMP Server Install Process

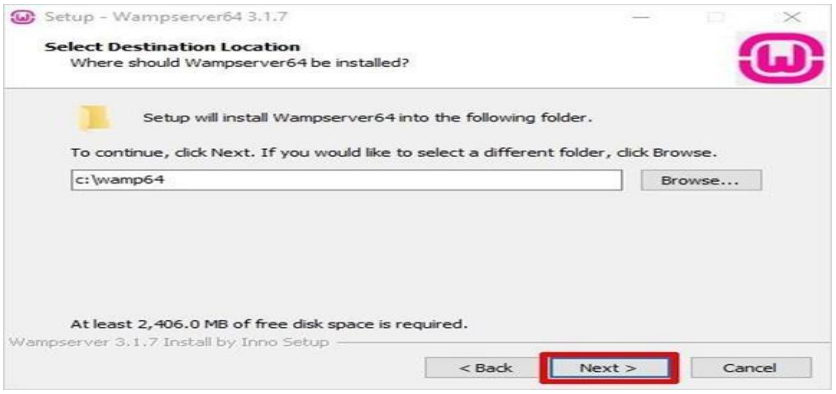
Soon after you click on the downloaded file, you will bask to choose your preferredlanguage and click the "Ok" button. As I'm choosing English here. See below.



The next screen will inform you about the required components that have to be present in your computer system for the proper functioning of the Wamp Server. The main purpose of this step is to check for the suitable version of Microsoft VC++ redistributable package.Click on the 'Next' to continue.

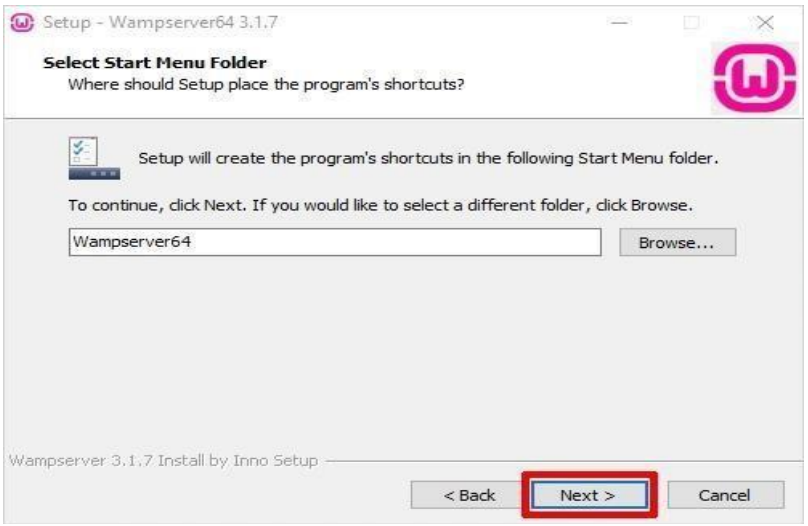
Step 3: Select Location/Destination to Install WAMP

"Here, you need to select the location to install the WAMP on your computer. However, you can accept the default location and move ahead. To continue with the installation hit the 'Next' button."



Step 4: Select Start Menu Folder to Install WAMP

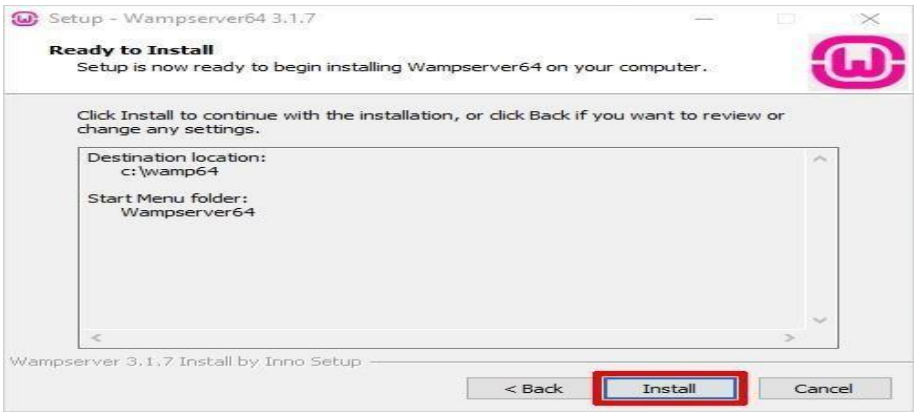
"Now, you need to select the folder where you wish to create the program's shortcut. You can select any folder of your choice or you can continue with the default option. You can easily access this folder from the Windows start button. Choose the folder andclick 'Next' to continue."

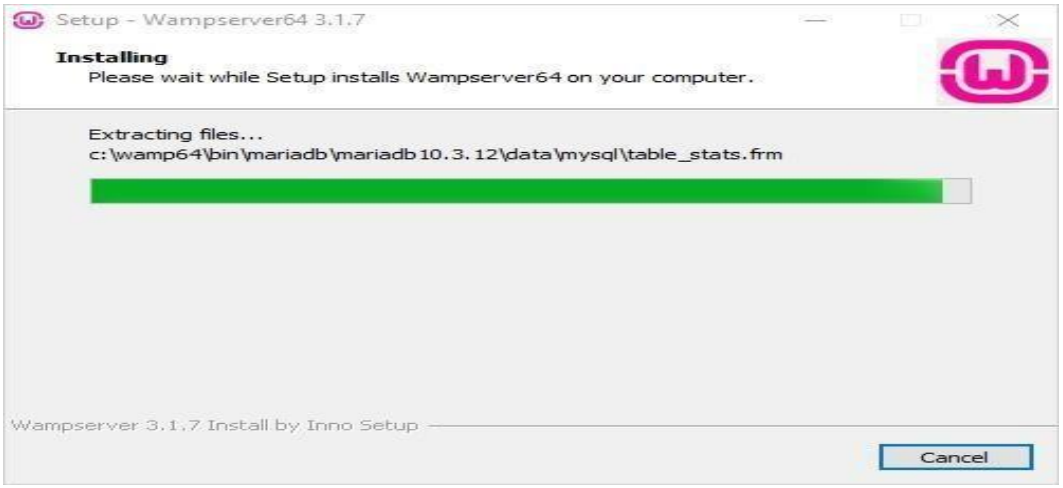


Step 5: Ready to Install WAMP

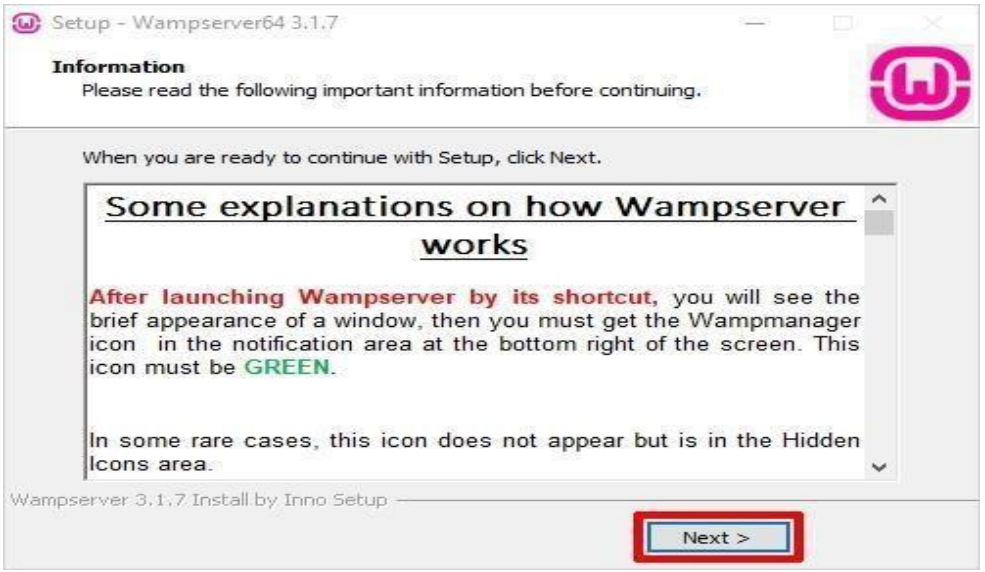
Finally, you would see a ready to install wizard with are add set up to begin theinstallation process.

Now, just is patient as the Wamp Server is extracting files to your selected location. Let the process completed.

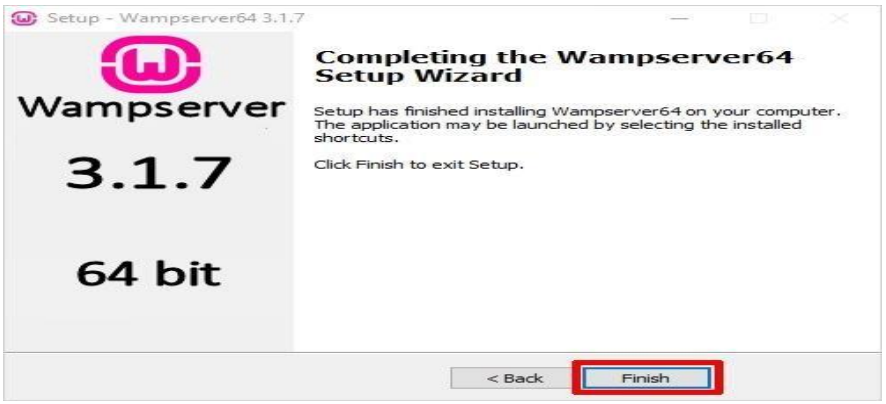




Once status bar is turned fully green, an information screen will appear like the one shownbelow. Click next to continue.



Step 6: WAMP Installation complete
The installation is almost done now simply click finish to exit setup.



REFERENCES

[1]. Arun Kumar, Naveen Kumar, Vishal Vats, "Efficient Crop Yield Prediction Using Machine Learning Algorithms", International Research Journal of Engineering and Technology(IRJET)-e-ISSN:2395-0056, p-ISSN:2395-0072, Volume: 05 Issue: 06, June-2018

[2]. Nithin Singh & saurabh chaturvedi, "Weather Forecasting Using Machine Learning", International Conference on Signal Processing and Communication (ICSC) Volume: 05, DEC-2019.

[3]. Aakash Parmar & Mithila Sompura, "Rainfall Prediction using Machine Learning", International Conference on (ICIIECS) at Coimbatore Volume: 3, March 2017.

[4]. Sachee Nene & Priya, R "Prediction of Crop yield using Machine learning", International Research Journal of Engineering and Technology (IRJET) Volume: 05 Issue: 02, Feb-2018.

[5]. Ramesh Medar & Anand M. Ambekar, "Sugarcane Crop prediction Using Supervised Machine Learning" published in International Journal of Intelligent Systems and Applications Volume: 3, August 2019.

[6]. Andrew Crane Droesch, "Machine learning methods for crop yield prediction and climate change impact assessment in agriculture", IOP Publishing Ltd Volume:05, OCT -2018.

[7]. Vinita Shah & Prachi Shah, "Groundnut Prediction Using Machine Learning Techniques ", IJSRCSEIT. UGC Journal No: 64718, March-2020.

[8]. Renuka & Sujata Terdal, "Evaluation of Machine Learning Algorithms for Crop Prediction" Published in International Journal of Engineering and Advanced Technology (IJEAT) Volume-8, August 2019.

[9]. P. Vinciya, Dr. A. Valarmathi, "Agriculture Analysis for Next Generation High Tech Farming in Data Mining" IJARCSSE, Volume. 6, Issue 5, 2016

[10]. S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika and J. Nisha, "Crop recommendation system for precision agriculture", IOP Publishing Ltd Volume: 05, 2014.

[11]. R. Kumar, M. P. Singh, P. Kumar and J. P. Singh, "Crop Selection Method to maximize crop yield rate using machine learning technique", International Journal of Engineering and Advanced Technology (IJEAT) Volume-8, 2016.

[12]. T.R. Lekhaa, "Efficient Crop Yield and Pesticide Prediction for Improving Agricultural Economy using Data Mining Techniques", IOP Publishing Ltd Volume: 05, 2020.

[13]. Viviliya, B. and Vaidhehi, V., "The Design of Hybrid Crop Recommendation System using Machine Learning Algorithms", International Journal of Engineering and Advanced Technology Volume-8, 2014.

[14]. Dahikar, S. S, Rode and S. V., "Agricultural crop yield prediction using artificial neural network approach", International Journal of Engineering and Advanced Technology (IJEAT) Volume-8, 2014.

[15]. Gonzlez Snchez. A, Frausto Sols. J and Ojeda Bustamante.W, "Predictive ability of machine learning methods for massive crop yield prediction", IOP Publishing Ltd Volume: 05, 2019.