A Project Report

on

Machine Learning in Agriculture: A New Farm-Hand

Submitted in partial fulfillment of the requirement for

the award ofthe degree of

Bachelor Of Technology
(Computer Science)

Under The Supervision

Ms. Kiran Singh:

Assistant Professor

Submitted By

Kunal Paswan
18021011517/18SCSE1010280

Garv Chaudhary

18021011714/18SCSE1010486

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA INDIA
10/2021

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING GALGOTIAS UNIVERSITY, GREATER NOIDA

CANDIDATE'S DECLARATION

I/We hereby certify that the work which is being presented in the thesis/project/dissertation, entitled "**Machine Learning in Agriculture: A New Farm-Hand**" in partial fulfillment of the requirements for the award of the Bachelor of Technology submitted in the School of Computing Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of September, 2021 to December,2021, under the supervision of Dr Kiran Singh, Department of Computer Science and Engineering/Computer Application and Information and Science, of School of Computing Science and Engineering , Galgotias University, Greater Noida.

The matter presented in the thesis/project/dissertation has not been submitted by me/us for the award of any other degree of this or any other places.

Kunal Paswan 18SCSE1010280
Garv Chaudhary 18SCSE1010486

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Supervisor
Dr. Kiran Singh

CERTIFICATE

The Final Thesis/Project/ Dissertation Viva-Voce examination of Kunal Paswan: 18SCSE1010280 and Garv Chaudhary: 18SCSE1010486been held on and his/her work is recommended for the award of Bachelor of Technology.

Signature of Examiner(s) **Signature of Supervisor(s)**

Signature of Project Coordinator **Signature of Dean**

Date: December, 2021
Place: Greater Noida

Acknowledgement

We would wish to precise our special thanks of gratitude to our major project guide professor "**Dr. Kiran Singh**" for his valuable guidance in completion of our project. We are also grateful to our Reviewer "**Dr. KM Balamurugan**" who supported us by through ideas to reinforce the quality and accessibility of our site. We are also thankful to our department of computing and technology for providing us with all the facilities required to end this project of ours.

Abstract

Growth in the agriculture domain is one of the major concerns of our country as it has a very impactful contribution in our country's economy. There are many problems that Indian farmers faced and those are the reasons why Indian farmer is so much behind. Those problems include problems like inequality in land distribution, small fragment land holding or lend tenure system and many others like that. Poor farming techniques are also one of the main reasons behind it. Various schemes are introduced by government to help the farmers but there are hardly any which can guide farmer in step by step manner. Survey used for the betterment and the educational purpose will have the major use of our system.

Our System is a Web based recommendation system in which process focuses on techniques like data mining for identifying the appropriate fertilizer. Our system uses the sufficiency approach which is used to check the nutritional value and approaches the best possible combination according to the crop requirement and the combination with lowest cost of fertilization. Our current dataset contain data about some of our major crops, temperature information and the soil quality information.

India is a nation where agriculture plays a very important role in country's economy. Prosperity of farmers means prosperity of nation. Thus, our work would help farmers in choosing the appropriate fertilizers which will improve their final yield. Our future scope will be aimed at an improve dataset with a larger number of attributes and also implement crop recommendation. Same application in future can be linked to some other facilities to provide other benefits for farmer to make it as a one solution for all farmer problems.

Table of Content

1. Introduction

1.1 Introduction

From Ancient time, Agriculture is considered as a very important practice in India and even now, Farming is still one of the major supports of all of the economic growth. It serves as much as quarter of Indian economy, employing more than 40 percent of Indian labour force[12]. Despite of constant progress in Agriculture Domain, the farmers keep facing problem which is very frustrating. Declining in the investment in agriculture domain is one of the reasons that would explain the decrease in growth rate of agriculture. Thus, there need of some methodologies and technologies to work together and came up with a system to help farmers[13].

To overcome the problem faced by farmers, Government had proposed and implemented many welfare schemes which help farmers in financial, marketing, irrigation, crop insurance and subsidies sector but the problem in this is the lack of exposure[11].

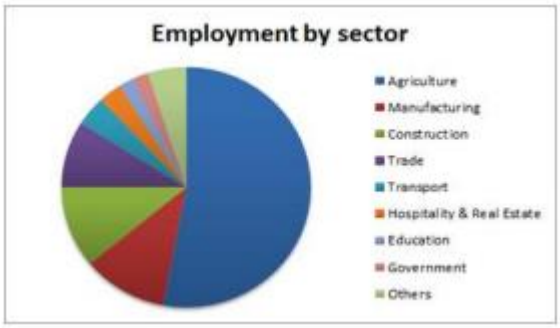
To work with this, we propose a system that would help farmers by recommending them best fertilizers for a specific crop and the soil quality. Proposed System is based on agriculture and revolves around the problem faced by many farmers all over the world. In this System, Users only need to provide the name of the crop, and our application automatically provides the data for the best recommended fertilizer available and then the alternatives if the best there is not available. It will also provide the right usage details and correct price for which it should be available in market[17]. This will improve the quality of farming and since this all is just one click process on your mobile, it will be time efficient as well. The system shows experimental outcomes as a section of implementation of proposed architecture. The farmer types query (crop name) to the search engine, in order to get information for the fertilizer of specific crop. Query asked are matched to fertilizer data, and then query is transformed into a semantic web query and after performing reasoning and semantic processing, the result is sent to the mobile device of user[18][19].

1.2 Formulation of Problem

Over a period of time Agriculture had faced many problems which challenges the agriculture to meet the demand of the need of the future. Now a days three main problems that pressurize agriculture are demographic, scarcity of natural resources and the climate change[15]. Except that it is also estimated that about one fourth of India's Production is lost due to the use of improper methodologies of farming. There are so many problems like lack of mechanization, small and fragmented land holdings, land tenure system, instability and fluctuations that were faced by farmers while choosing the right fertilizers for a crop, matching the right amount of fertilizer for that crop, and getting those fertilizers at the correct price from the market. Fertilizer problems may not sound too troublesome and get unnoticed sometimes but it can have a major effect on crop final yield and in some cases like in the case of over fertilization it can affect the soil which will be a huge trouble in the long run. So, using right fertilizers and then applying them in proper amount for the right time of the year is necessary as it can enhance the final yield of crop and can help farmers to get better crop production[16].

1.3 Statistics

Agriculture is one of the benchmark areas and milestone feelings of mankind. The word agriculture is found in the Latin language where ager is land and culture is cultivation. According to the Food and Agricultural Organization of UN the humans depend on agriculture for their survival more than sixty percent and around twelve percent of the land is for the crop production. Presumed by the Food and Agricultural Organization, the whole native would make it eight billion by the year 2025 then ten billion by the end of 2050[21]. To manage this growth in population an increase of seventy percent in crop production is estimated to achieve all over the globe by 2050. Agriculture has been ranked 2nd for emitting greenhouse, biomass, machinery and fossil-based fertilizers and also it is the backbone of the economy for some of the nation. Figure below shows the growth of the employment sector all over the world.



Employment by sector with GDP Growth (2016-2021)

Countries	5 year Moving Average to 2016	20 year Moving Average to 2016	IMF Projected Annual Growth 2016-2021
		BRICS	
Brazil	-0.40%	2.30%	1.60%
Russia	0.50%	3.30%	1.50%
India	6.80%	6.90%	7.70%
China	7.30%	9.20%	6.10%
South Africa	1.60%	2.80%	1.80%
		G7	
US	2.10%	2.30%	2.10%
Japan	1.20%	0.70%	0.70%
Germany	1.20%	1.40%	1.80%
UK	2.10%	2.00%	1.40%
Canada	1.80%	1.90%	1.90%

Table. GDP Growth in BRICS & G7

In India fifty three percent of the nation's natives are totally relying on agriculture for jobs and sixty two percent of the nation's natives are dependent essentially on agriculture for their survival[24]. By Focusing on the size of the market, India is the 2nd largest in producing the fruit industry across the globe. Based on the statistics, the income of farming in India is to be expected to double by the year 2022. Due to the increase in the demand with the rise of population, results are unstable in the coming years. Some of the important factors that should be taken into consideration are global warming and climate conditions. According to the contributions of Machine Learning and its related technologies in the development and advancement of agricultural fields, this research has been done by going through some of the articles.

1.4 Tools and Technologies Used

In Fertilizer recommendation system, the process focuses on the use of data mining techniques for identifying appropriate fertilizer and then recommending them. Our system has a sufficiency approach which is used to check the nutritional value and approach the possible combination of fertilizer for the crop requirement and combination with the lowest cost of fertilization will be recommended.

Tools Used:

i) **Anaconda Navigator:** Anaconda Navigator is a graphical user interface which includes Anaconda distribution that helps in the launching of application and in the management of conda packages, environment and channels without using the command line system. It is a package manager, an environment manager, a Python/R data science distribution, and a collection of over 7,500+ open-source packages. It is available for operating systems like Windows, macOS and Linux.

ii) **Jupyter Notebook:** The Jupyter Notebook is an open -source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. Jupyter Notebook is maintained by the people at Project Jupyter. It can be used for data science, statistical modeling, machine learning and many such things.

iii) **MySql:** It is a relational database that organizes data into one or more data tables in which data types may be related to each other; these relations help structure the data. SQL is a language programmers use to create, modify and extract data from the relational database, as well as control user access to the database.

iv) **Visual Studio:** Visual Studio is an integrated development environment created by Microsoft for the development of computer programs. It also includes the development of websites, web application and mobile applications. It includes built-in languages like C, C++, C++/CLI, Basic.NET, C#, F#, JavaScript, TypeScript, XML, XSLT, HTML and CSS. It also supports languages such as Python, Ruby, Node.js and M among others.

Technologies Used:

i) **Django:** Django is a high-level Python web framework that enables rapid development of secure and maintainable websites. Built by experienced developers, Django 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 is free and open source, has a thriving and active community, great documentation, and many options for free and paid-for support
.

ii) **Machine Learning:** Machine Learning System automatically learn program from data. It used in web Search, spam filter, recommendation system, and placement, credit scoring, recommendation of stock trading, fraud detection etc. There are lot of types of Machine learning but the one that used in our system is:

a) **Classifier Classification:** Input is a vector of discrete or continuous form and Output is a single discrete value. This kind of Machine learning is used for predicting and recommending.

Our final model will be judge on the basis of Accuracy(percentage of correct prediction), Precision(defined as the fraction of relevant example among all of examples predicted to belong in a certain class) and Recall(defined as fractions of example predicted to belong to a class with respect to all of examples that truly belongs in class).

iii) **Front End:**

a) **HTML**: The hyper text markup language or HTML is the standard markup language for documents designed to be displayed in the web browser. It can be assisted by technologies such as cascading style sheets(CSS) , and scripting languages such as JavaScript.

b) **CSS:** We use CSS for the designing of the web application. Cascading Style Sheets (CSS) is a style sheet language use for describing the presentation of the document written in the markup language such as HTML. CSS is a cornerstone technology of the world wide web, alongside HTML and Javascript. CSS is designed to enable the separation of presentation and content including layout, colors and fonts.

iv) Back End:

a) Python: It is a high-level general-purpose programming language. It is dynamically typed and also is garbage-collected. Multiple Programming and object-oriented and functional programming are one of the main features of python programming. Flexibility in Python allows it to be a great option for Machine Learning. Developing Scripts in Python is much easier because of all the standard libraires that are present in Python.

b) SQL: SQL stands for Structured Query Language. It is the standardize language that is used to manage the relational database. It allows users to perform various operations on the database and the retrieve data according. It also includes adding, updating, removing, creating and retrieving only some particular information of data from database too. It used SQL syntax for its coding format and writing statement.

2. Literature Survey

This section comprises of the finding related to our topic. It is basically a report on the information we found in literature relevant to fertilizer, soil, crop and nutritional value. The system existing are crop recommendation system, Query analysis and recommendation with Hadoop framework. Some of the many reported domain walks in this system are:

- **Efficient Crop Yield Recommendation by the Machine Learning for Digital Farming**

Author: Dr. G Suresh et al [1]

This paper proposes that how the efficient crop yield can be recommended by the machine learning in digital farming. Different Machine Learning algorithms like Logistic regression, random forest algorithm, support vector machine algorithm was tested on training data and then compared on the basis of their prediction accuracy.

- **Agriculture Market Recommendation System**

Author: K.Dhivyaa et al [2]

This paper proposes ways for the crop recommendation on the basis of the soil condition. In this paper, different attributes of the soil is analyzed and a machine learning model was trained to predict the crop recommendation. Paper also has little information about market pricing and fertilizer recommendation.

- **Crop Recommendation System to Maximize the Crop yield in Ramtek Region Using Machine Learning**

Author: Dr. Ananatha Reddy et al [3]

This paper aims at Precision Farming. Paper tells about how factors like choosing the right crop according to soil conditions and the choosing the appropriate farming techniques like fertilizer and pesticide can help in achieving the maximize crop yield.

- **A Survey on Machine Learning and Text Processing for Pesticides and Fertilizer Recommendation**

Author: P. Pandiraja et al [4]

This Paper Survey Machine Learning Methodologies along with text processing System in agriculture. Marking the future Challenging and the situation of current problems in agriculture field are thoroughly assessed in this paper. Paper survey's main focus was the random forest algorithm and flutter framework in Machine Learning.

- **Data Mining Technique to Analyse soil Nutrition based on Hybrid Classification**

Author: E.Manjula et al [5]

This paper chooses Nitrogen, Phosphorus, Potassium, Calcium, Magnesium, Sulphur, Iron, Zinc, and so forth, nutrients for investigating the soil supplements utilizing Naïve Bayes, Decision Tree and hybrid approach of Naïve Bayes and Decision Tree. The performances of the classification algorithms are compared based on accuracy and execution time.

- **Prediction of crop yield and fertilizer Recommendation System Using Machine Learning Algorithms:**
  Author: Devdutta A Bondre et al [6]

  This paper comprises of a system that helps with the prediction of crop yield and fertilizer both. Fertilizer was recommended in this system was on the basis of crop type and location. There were some pre-specified fertilizers for few particular sets of crops that were recommended on the basis of the location. Some Machine Learning Algorithms were also get compared to get to know that which among them would be best for achieving their object of yield prediction.

- **Collaborative Recommendation System for Agriculture Sector:**
  Author: Sapna Jaiswal et al [7]

  This paper develops a research system that help farmer with their general queries by rectifying them in minutes or second through a web application. This includes the generation of their profile and then updating them with new trends and new government schemes according to the data they gave during the time of profile generation.

- **Agriculture Price Prediction Using Data Mining:**
  Author: Veeresh Kadlimatti, S V Saboji et al [8]

  Cultivating is the basic source and spine of Indian Economy and expects to be a basic part in individual life. In indisputably the GDP (Gross domestic product) cultivating nearly contributes sixteen percent and for extending new exchange it contributes practically 10% to the total country conveys. As the general population reliably extending and to manage crafted by the country there's requires a genuine use and the leaders of cultivation things. Data mining is a predominant methodology and most ideal choice in expecting the exact expenses of the cultivation reliant upon past data. In this work distinctive data mining computations are applied on the dataset to expect what's to come expenses of the agribusiness things.

- **Analysis of crop yield prediction using data mining technique to predict annual yield of major crops.**
  Author: B. Devika, B. Ananthi et al [9]

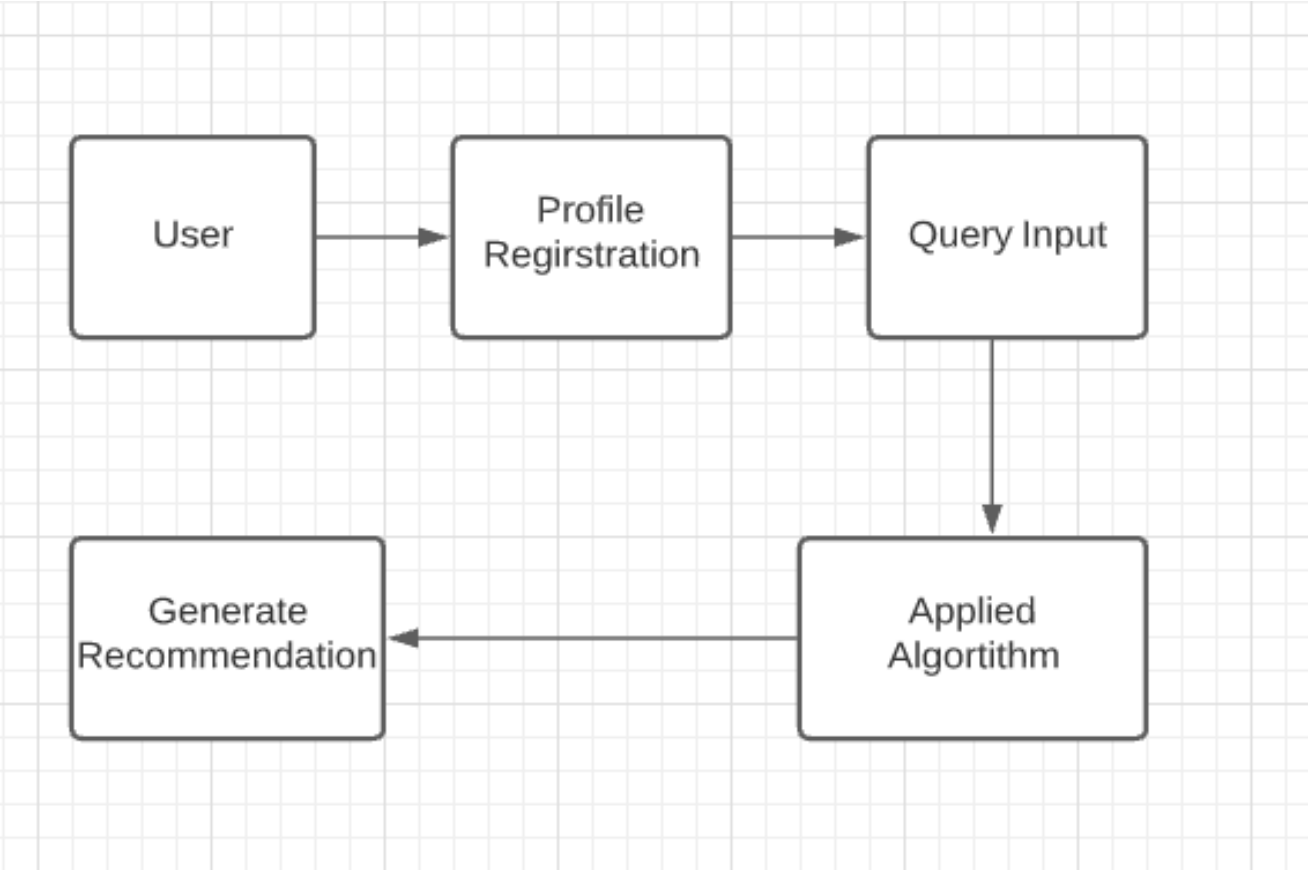
  India is generally a plant country. Agribusiness is irrefutably the main provider to the Indian economy. Cultivating harvest creation depends upon the season, normal, and monetary explanation. The speculation of agrarian yield is trying and fulfilling task for every country. Nowadays, Ranchers are taking steps to make the yield because of conflicting climatic changes and lack of water resource. The essential objective is gathering plant data which can be taken care of and separated for supportive gather yield deciding. To anticipate the reap yield with the help of data mining methodology, advanced systems can be familiar with expect crop yield and it similarly supports the farmer to pick the most fitting harvest, as needs be working on the value and gain of the developing region.

- **Existing System:**

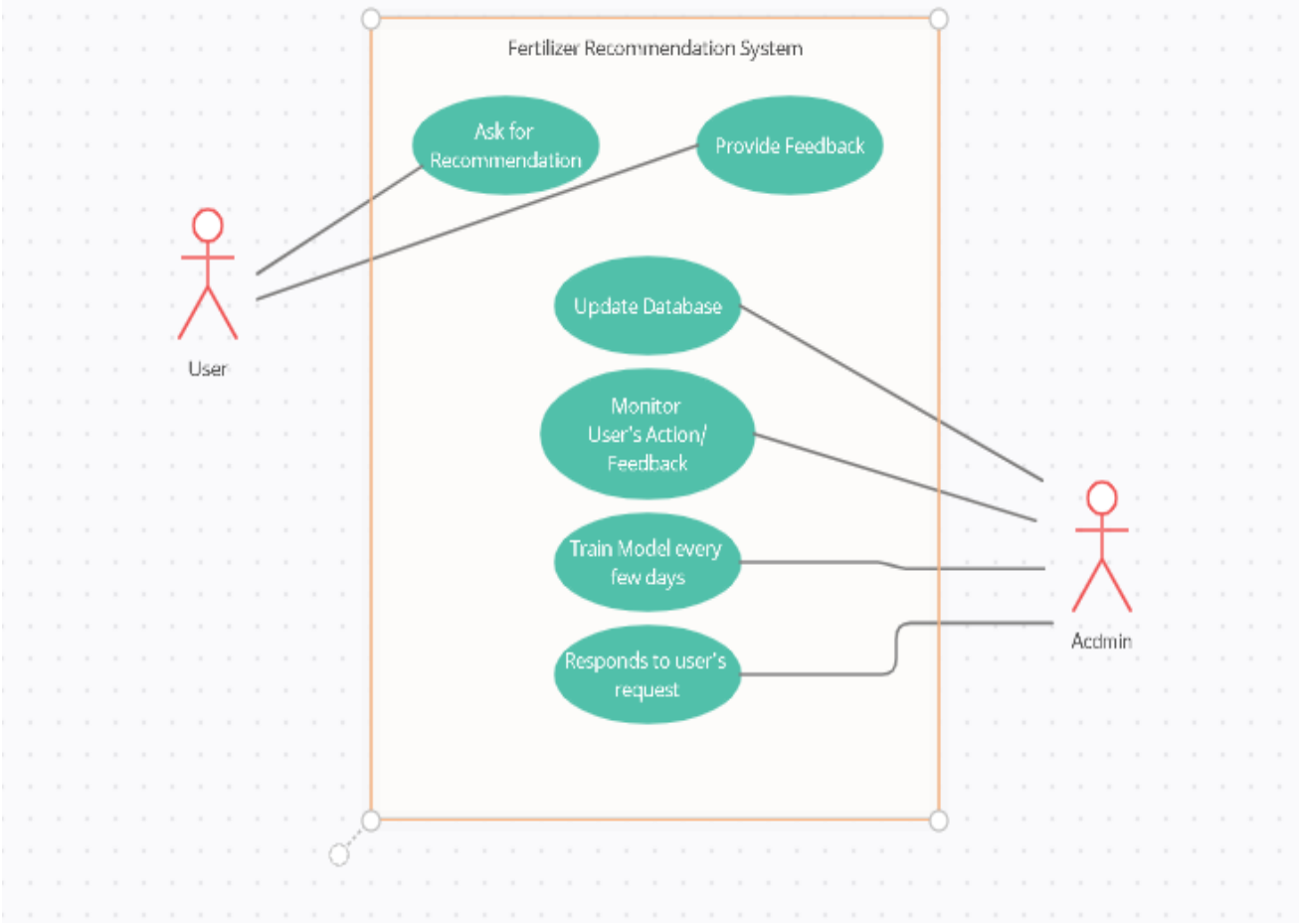
  Work done till now mainly focuses on the crop prediction using soil quality and data mining technique. Fertilizer Recommendation is note taken into consideration solely. There are some works that gave little hint about our system but there was no implementation available for that also. Machine Learning algorithms like Naïve Bayes, Random Forest Algorithm, Decision Tree or some other Algorithms too were part of previous System but none was used for Fertilizer Recommendation System[10]. So, our system will develop a fertilizer recommendation system that will predict the fertilizer on the basis of two things: first is the crop type and soil type in which simple data visualization process was used and second is the process in which we develop a machine learning model with the most suitable algorithms that predicts the best fertilizer on the basis of the crop nutritional requirement[14].

3. Project Design

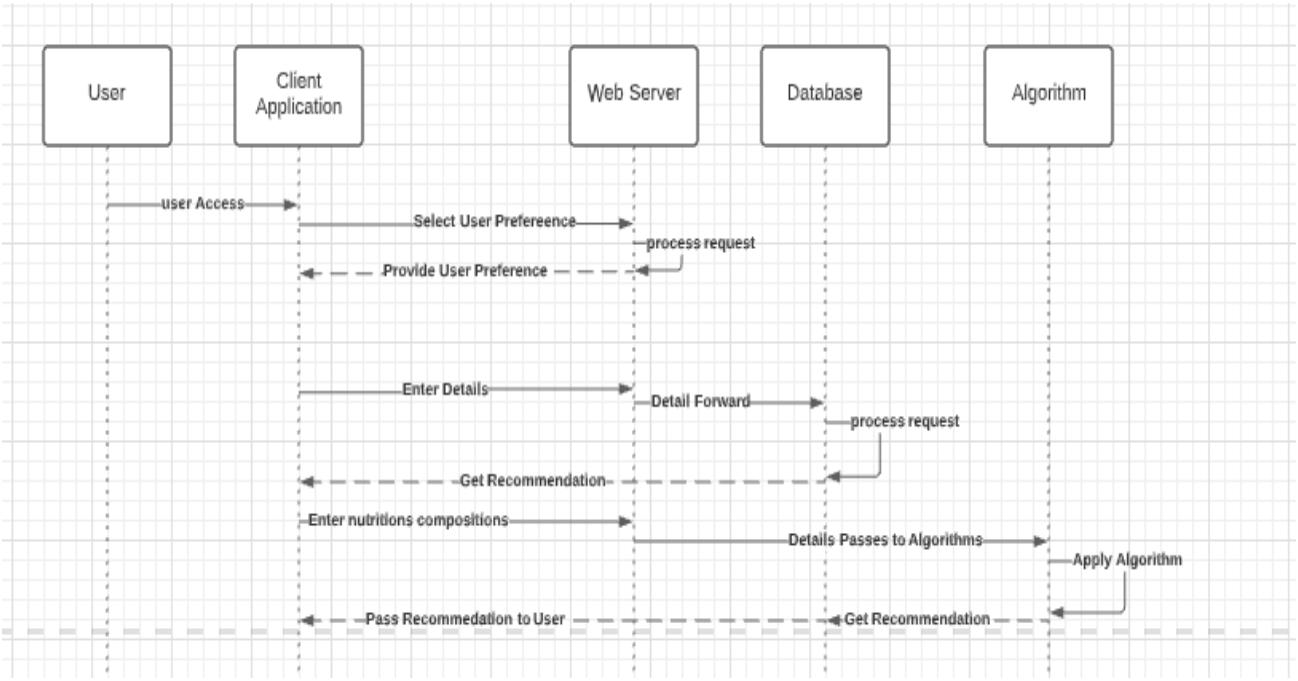
3.1 Proposed System:



3.2 Use case Diagram:



3.3 Sequence Diagram



3.4 Database:

https://www.kaggle.com/gdabhishek/fertilizer-prediction

4. Module Description

Since we used Django Framework. All the modules that are designed are basically used languages like Python, HTML or CSS. There is a homepage which provide two options to the user and on the basis of which option user may picked, he/she may get redirected to the next page. Next redirected pages consist of the project implementations and provide further options to user to enter the required information.

4.1 Front End Development:

- **Homepage:** So below is the source code of Homepage and how homepage looks:

```
 <!DOCTYPE html>
<html lang="en">
<head>
        <meta charset="utf-8">
        <meta name="viewport" content="width=device-width, initial-scale=1, shrink-to-fit=no">
        <meta name="description" content="">
        <meta name="author" content="">
        <title>Fertilizer Recommendation System </title>
   <style>
      a:link {
         color: black;
         background-color: transparent;
         text-decoration: none;
      }
      a:visited {0-
         color: black;
         background-color: transparent;
         text-decoration: none;
      }

      a:hover {
         color: red;
         background-color: transparent;
         text-decoration: underline;
      }

      body {
         background-image: url("{% static 'images/dawn.jpg' %}");
         background-repeat: no-repeat;
         background-attachment: fixed;
         background-position: center;
      }
   </style>
        <!-- Bootstrap core CSS -->
        <link href="{% static 'vendor/bootstrap/css/bootstrap.min.css' %}">
        <!-- Fontawesome CSS -->
        <link href="{% static 'css/all.css' %}">
        <!-- Custom styles for this template -->
```

```
        <link href="{% static 'css/style.css' %}">
</head>
<body>
   <!-- Slide One - Set the background image for this slide in the line below -->
      <h1>Fertilizer Recommendation System</h1>
      <p><h4>Here, we recommend you the fertilizer which would be best for your crop.<br> You can get recommendation based on Crop type and Soil type Or You can get recommendation based on the Nutritional Requirement of you Crop.</h4></p>

   <!-- Slide Three - Set the background image for this slide in the line below -->
         <a href="click"><h2>Click here if you want recommendation based on Crop type and Soil type</h2></a>
         <a href="click1"><h2> Click here if you want recommendation based on the Nutritional Requirement of your Crop</h2></a>

            <!-- /.container -->
            <!--footer starts from here-->
            <footer class="footer">
               <div class="container bottom_border">
                  <div class="row">
                     <div class="col-lg-3 col-md-6 col-sm-6 col">
                        <h5 class="headin5_amrc col_white_amrc pt2">Contact us:</h5>
                        <p><i class="fa fa-phone"></i> +91 - 1234567890</p>
                        <p><i class="fa fa fa-envelope"></i> kp@example.com </p>
                     </div>
                  </div>
                  </div>
               </footer>

</body>
            </html>
```

Homepage of Web Application:

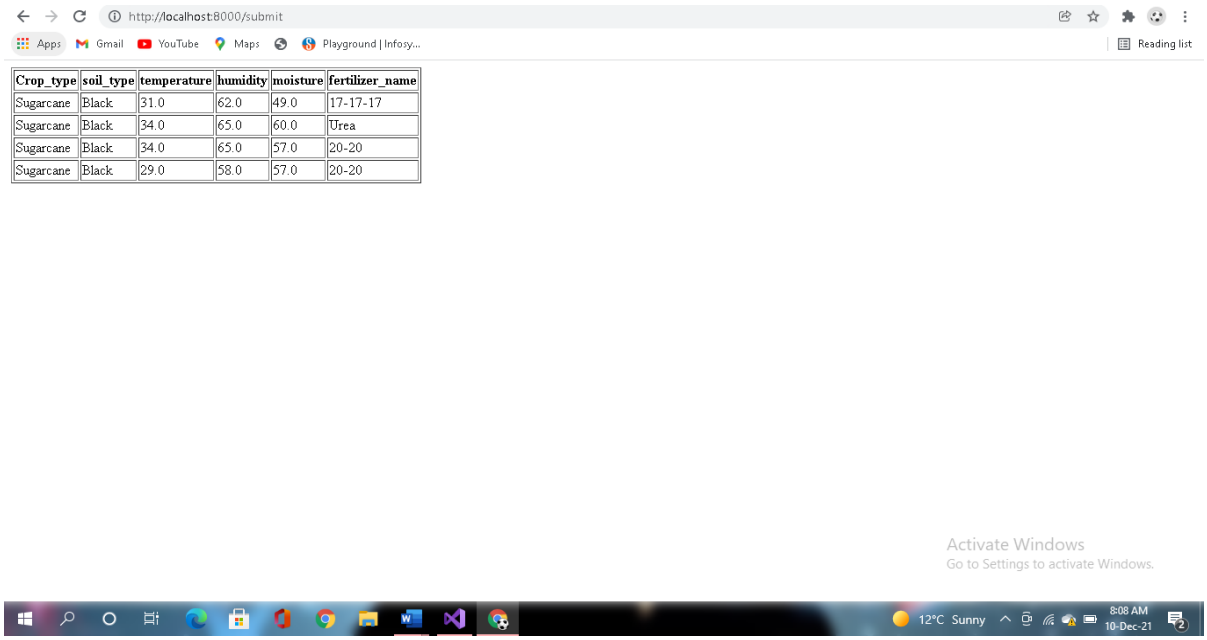
There are two options in the home page that can be clearly seen in the image above. First one is ‘click here if you want recommendation based on Crop Type and Soil Type’ and the other one is ‘click here if you want recommendation based on the Nutritional Requirement of your crop’. You will be redirected to next page according to the option you choose.

- **Source code for the page that user get redirected on selecting option 1:** If you choose option 1 then it will redirect you to the page that ask you to select the crop name and soil type from a drop-down menu. You choose the combination of both and then can press the submit button for submitting your query.

```
  <h2> <label for=" crop">Enter Crop name:</label>
   <select name="crop" id="crop">
      <option>Sugarcane</option>
      <option>Millets</option>
      <option>Maize</option>
      <option>Paddy</option>
      <option>Wheat</option>
      <option>Barley</option>
      <option>Oil seed</option>
      <option>Tobacco</option>
      <option>Pulses</option>
      <option>Ground Nuts</option>
      <option>Cotton</option>
   </select><br></h2>
   <h2><label for="soil">Enter Soil type:</label>
   <select name="soil" id="soil">
      <option>Black</option>
      <option>Loamy</option>
      <option>Red</option>
      <option>Clayey</option>
      <option>Sandy</option>
   </select><br></h2>
   <input type="submit">
</form>
```

The page that user get redirected to if he/she select option1:

As it can be seen, there are spaces where you select your preferred types. In the above image, the selected Crop type is Sugarcane and the soil type is black. So, this information is passed forward to our system when the submit button is pressed and our system after processing that, will predict the best fertilizers according to the location's temperature and humidity. Below is the output that we get by processing above data:



Crop_type	soil_type	temperature	humidity	moisture	fertilizer_name
Sugarcane	Black	31.0	62.0	49.0	17-17-17
Sugarcane	Black	34.0	65.0	60.0	Urea
Sugarcane	Black	34.0	65.0	57.0	20-20
Sugarcane	Black	29.0	58.0	57.0	20-20

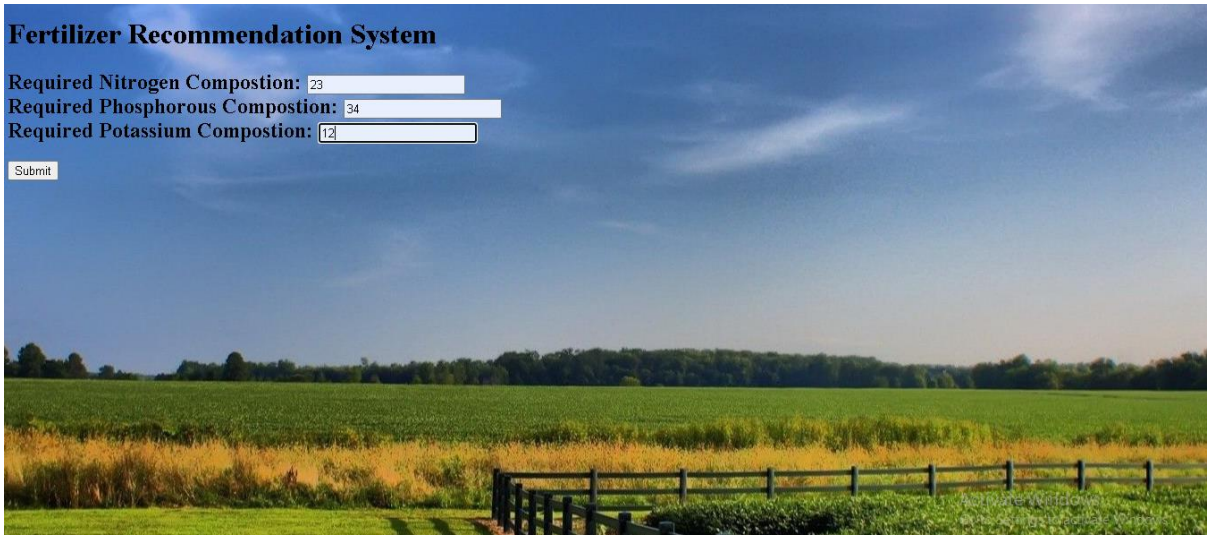
So, in above image it can be seen that all the data from our current database which have same crop type and soil type had been retrieved and shown on screen. Different fertilizers are recommended on the basis of location's temperature, moisture and humidity helping to choose best among best fertilizers out there.

- **Source code for the page that appears on selecting option 2:** When user select this option, he must know the nutritional requirement of the crop they be planting in composition. This information is available on the internet and user can get that from there. After selecting this option, User get redirected to a page which ask for three major nutritional composition, that user have to provide. That information is passed to our model which takes the data and process in the order in which it is trained and then recommend the fertilizer.

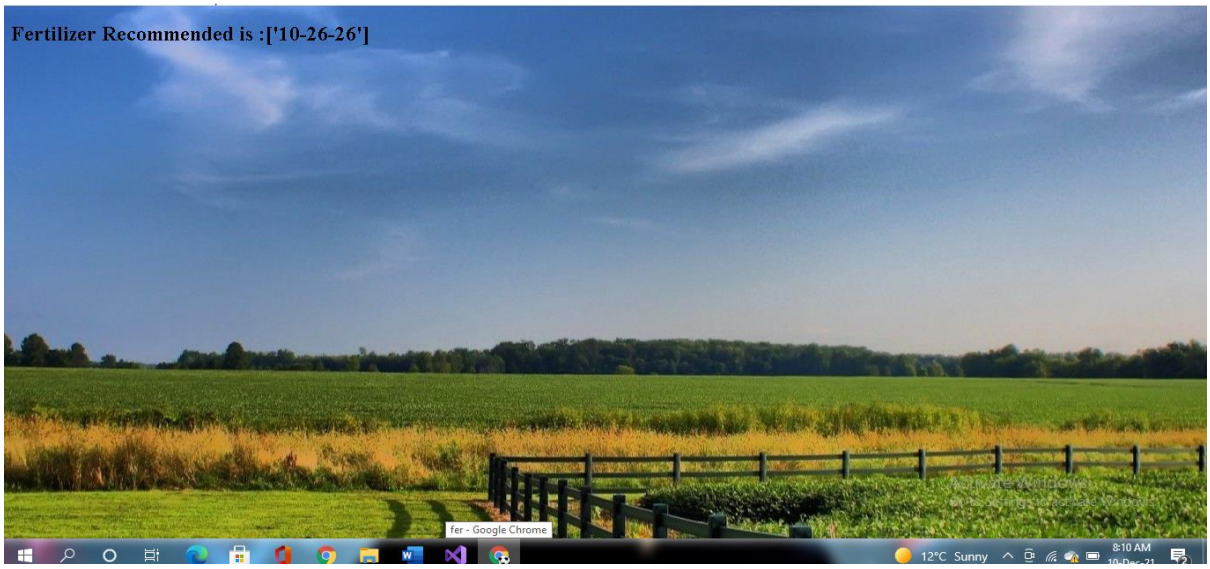
```
<h1>Fertilizer Recommendation System</h1>
<form action="predict" method='post'>

    {% csrf_token %}
    <h2>
    Required Nitrogen    Compostion: <input type="text" name='Nitro'><br>
    Required Phosphorous Compostion: <input type="text" name='Phos'><br>
    Required Potassium   Compostion: <input type="text" name='Potas'><br>
      </h2>
    <input type="submit">
```

The page that user get redirected to if he/ she choose option 2:



In the above page, it can be seen that it clearly asked for the three major nutritional composition that is NPK (Nitrogen, Phosphorous and Potassium) of the crop user planting. The information entered in above image is 23 percent for the Nitrogen, 34 percent for Phosphorous and 12 percent for the potassium. On pressing the submission icon, this information is passed to our model which will process that data and will recommend you the fertilizer. On pressing the icon, below is the page that you will get redirected to



So after taking all the inputs and processing, our model recommend users to get the fertilizer named with ’10-26-26’.

4.2 Back End Development

For this purpose, we use Python as it was clearly stated before too. Mysql was used for for database purpose and then it was connected to python for further implications.
In python have use several libraries, and information about few of them that are used to develop this system is below:

- **Numpy:**

  It offers powerful N-Dimensional Array that are fast and versatile and also helps vectorization and indexing. It also offers itself as a numerical computing tool which can solve many mathematical functions, linear algebra routines or fourier transformations. The core of Numpy is a well-optimized C code so it provides flexibility with the speed.

- **Panda**

  Panda is the fast easy to use tool which is use for data analysis that have been built over the python programming language. It is a powerful tool which can be used for data manipulation. It is one of the most important libraries in the field of Data Analysis and Data Science.

- **Mysql_connector**

  Python need driver to access Mysql servers. This is that self-contained Python driver which is used for communication with Mysql servers and helps in developments of database applications. There are few steps to create the connection between database and python and after those easy steps, connections formed helping in building of those database applications.

- **Sklearn:**

  It is a simple and efficient tool which is used for data prediction. It is open source so it is available to everyone and is reusable to various context. It is built on the basis of NumPy, SciPy and matplotlib. It can be used for classification, Regression and Clustering.

Code:

```
from django.shortcuts import render
from django.utils.datastructures import MultiValueDictKeyError
from django.http import HttpResponse
import mysql.connector
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn import svm

# Create your views here.
def index(request):
    return render(request,'index.html')
def click(request):
    return render(request,'home.html')
def click1(request):
    return render(request,'home1.html')
#option1
def submit(request):
    crop = request.POST.get('crop',False)
    soil = request.POST.get('soil',False)

    Connection= mysql.connector.connect(host= 'localhost',
                                        user='root',
                                        password='####',
                                        db='project',
                                            )
    b=crop
    c=soil
    r=pd.read_sql_query('Select Crop_type, soil_type, temperature, humidity, moisture, fertilizer_name from
fertilizer where crop_type= %(crop_type)s and
soil_type=%(soil_type)s',Connection,params={'crop_type':b,'soil_type':c})
    t=pd.DataFrame(r)
    t.to_html('C:/Users/Vishnu/projects/fer/template/tables.html',index=False)

    return render(request,'tables.html')
#option2
def predict(request):
    Nitro =int( request.POST.get('Nitro',False))
    Phos = int(request.POST.get('Phos',False))
    Potas = int(request.POST.get('Potas',False))
    k_test=[[Nitro, Phos, Potas]]
    fertilizer_df=pd.read_csv("C:/Users/Vishnu/projects/fer/template/Fertilizer Prediction.csv")
    fertilizer_df.columns
```

```
feature_df=fertilizer_df[['Nitrogen', 'Potassium', 'Phosphorous']]
X=np.asarray(feature_df)
y=np.asarray(fertilizer_df['Fertilizer Name'])
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,random_state=4)
classifier=svm.SVC(kernel='linear',gamma='auto',C=2)
classifier.fit(X_train,y_train)
y_predict=classifier.predict(k_test)

return render(request,'Res.html',{'result':y_predict})
```

4.3 Data Connectivity:

In the option 1, we use the concept of Data Visualization. We used a database and read all the information in the form of data frame and then used mysql connector to form the connection between sql servers and python. After establishing connection, we use sql queries to retrieve the data accordingly. Below are the code snippets of the program that was used during that process:

1.) Reading the data from database and determining its shape

```
In [1]: import pandas as pd
        fd=pd.read_csv("Fertilizer Prediction.csv")

In [2]: fd.shape

Out[2]: (99, 9)
```

2.) Importing all the required libraires and establishing connections

```
In [4]: import os
        import mysql.connector
        import pandas as pd
        import numpy as np
        %matplotlib inline

In [5]: connection=mysql.connector.connect(host='localhost',
                                           user='Kunal',
                                          password='####',
                                          db='project')

In [6]: connection

Out[6]: <mysql.connector.connection.MySQLConnection at 0x1f408204e88>
```

:

3.) Reading a query in the system and checking for its output:

```
b=input("Enter the crop name:")
fr=pd.read_sql_query('select Soil_type from fertilizer where crop_type= %(crop_type)s',connection,params={'crop_type':b})
fr
```

```
Enter the crop name:paddy
```

	Soil_type
0	Clayey

Connection was established and we can use Sql query in python to retrieve the data from the database as it can be clearly visible in above snippet.

4.4 Algorithms:

So, in the option 2 we used Machine learning. We trained a model on some data and then checked how it worked on test data. After getting satisfactory result, we used that model in our system. Our objective was to predict the best fertilizer that can be recommended to a certain crop which have some certain sets of nutritional requirements. So, we need to used to the algorithms that uses classifications[20]. There are several options available in Machine Learning that can be used for this purpose. Some of them are nearest neighbors, random forest and SVM (Support Vector Machine) etc.

1.) RANDOM TREE:

Random tree is similar to that of a decision tree but it differs from random tree in a way that for each split only a random subset of attribute is available. Random tree can be built for both nominal and numerical data. The Random tree is similar to C4.5 or CART butt it selects only a random subset of attributes. At each node, it considers K randomly chosen attributes. The subset ratio parameter specifies the size of the subset. Types of random trees include:

- Uniform spanning tree, a spanning tree of a given graph in which each different tree is equally likely to be selected
- Random minimal spanning tree, spanning trees of a graph formed by choosing random edge weights and using the minimum spanning tree for those weights
- Random binary tree, binary trees with a given number of nodes, formed by inserting the nodes in a random order or by selecting all possible trees uniformly at random
- Random recursive tree, increasingly labelled trees, which can be generated using a simple stochastic growth rule.
- Treap or randomized binary search tree, a data structure that uses random choices to simulate a random binary tree for non-random update sequences
- Rapidly exploring random tree, a fractal space-filling pattern used as a data structure for searching high-dimensional spaces
- Brownian tree, a fractal tree structure created by diffusion-limited aggregation processes

- Random forest, a machine-learning classifier based on choosing random subsets of variables for each tree and using the most frequent tree output as the overall classification
- Branching process, a model of a population in which each individual has a random number of children

2.) K-NEAREST NEIGHBOR:

The k-nearest neighbors (KNN) algorithm is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems. It's easy to implement and understand, but has a major drawback of becoming significantly slows as the size of that data in use grows.

KNN works by finding the distances between a query and all the examples in the data, selecting the specified number examples (K) closest to the query, then votes for the most frequent label (in the case of classification) or averages the labels (in the case of regression).

In the case of classification and regression, we saw that choosing the right K for our data is done by trying several Ks and picking the one that works best.

K-Nearest Neighbor can be used for both classification and regression. K-Nearest Neighbors is a non-complex algorithm which stores all the available cases and classifiers new cased based on some similarity measure. The sampled set is classified based upon the" closeness" that is the distance measure such as Euclidean distance or Manhattan distance.

It shows three operators namely retrieve, set role, validation. the retrieve operator retrieves the dataset that is being uploaded in the tool. the set role operator used to set the target attribute or special attributes. In order to estimate the statistical performance of learning operator a cross-validation is performed by the validation operator. The training process consist of the voting operator which is the technique that we process for the better results. On the testing sub-process lies the apply model and performance operator which evaluate the correctness of the model.

3.) RANDOM FOREST:

Random forest square measure associate ensemble learning methodology for classification, regression and different tasks, that operate by building a mess of call trees at coaching time and outputting the category that's the mode of categories or mean prediction of the individual trees. Random call forests correct for call tree custom of over fitting to them coaching set.
The primary rule for random call forest was created by Tin KamHo victimization the random mathematical space methodology, which, in Ho's formulation, could be a thanks to implement the "stochastic discrimination" approach to classification

A random forest is a machine learning technique that's used to solve regression and classification problems. It utilizes ensemble learning, which is a technique that combines many classifiers to provide solutions to complex problems.

A random forest algorithm consists of many decision trees. The 'forest' generated by the random forest algorithm is trained through bagging or bootstrap aggregating. Bagging is an ensemble meta-algorithm that improves the accuracy of machine learning algorithms.

The (random forest) algorithm establishes the outcome based on the predictions of the decision trees. It predicts by taking the average or mean of the output from various trees. Increasing the number of trees increases the precision of the outcome.

A random forest eradicates the limitations of a decision tree algorithm. It reduces the overfitting of datasets and increases precision. It generates predictions without requiring many configurations in packages (like scikit-learn).

4.) DECISION TREE:

Classifies data using the attributes. Tree consists of decision nodes and decision leafs. Nodes can have two or more branches which represents the value for the attributes tested. Leafs nodes produces a homogeneous result[22].

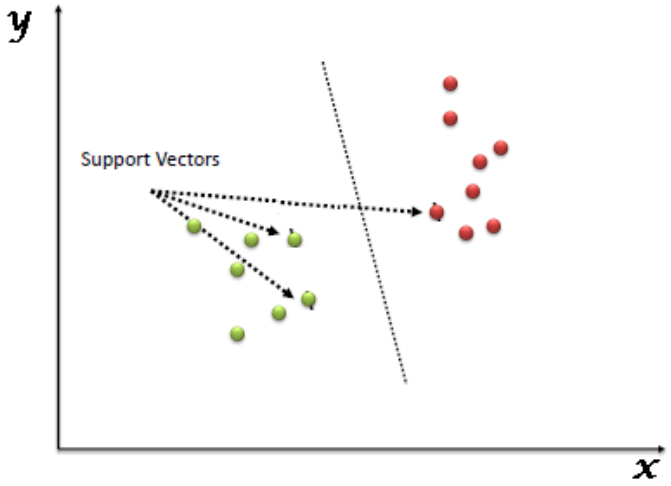
- Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.
- In a Decision tree, there are two nodes, which are the Decision Node and Leaf Node. Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches.
- The decisions or the test are performed on the basis of features of the given dataset.
- It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions.
- It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure.
- In order to build a tree, we use the CART algorithm, which stands for Classification and Regression Tree algorithm.
- A decision tree simply asks a question, and based on the answer (Yes/No), it further split the tree into subtrees.

4.4.1 Algorithms Used in System:

After reading several research papers, SVM is prove to be best algorithm for recommending fertilizers. So, we use SVM in our system for the predictions.

SUPPORT VECTOR MACHINE (SVM):

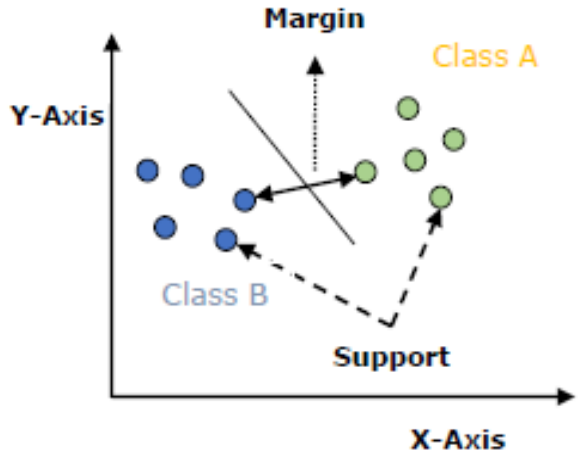
Support Vector Machine" (SVM) is a supervised machine learning algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is a number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well.



Support Vectors are simply the coordinates of individual observation. The SVM classifier is a frontier that best segregates the two classes (hyper-plane/ line)[23].

- **Advantages:**
  1) SVM calculation has a regularization parameter, which stays away from over-fitting.
  2) SVM calculation utilizes the portion trap, so you can construct master learning about the issue.

- **Working Of SVM:**
  An SVM model is basically a representation of different classes in a hyperplane in multidimensional space. The hyperplane will be generated in an iterative manner by SVM so that the error can be minimized. The goal of SVM is to divide the datasets into classes to find a maximum marginal hyperplane (MMH).



The followings are important concepts in SVM –

- Support Vectors − Datapoints that are closest to the hyperplane is called support vectors. Separating line will be defined with the help of these data points.
- Hyperplane − As we can see in the above diagram, it is a decision plane or space which is divided between a set of objects having different classes.
- Margin − It may be defined as the gap between two lines on the closet data points of different classes. It can be calculated as the perpendicular distance from the line to the support vectors. Large margin is considered as a good margin and small margin is considered as a bad margin.

The main goal of SVM is to divide the datasets into classes to find a maximum marginal hyperplane (MMH) and it can be done in the following two steps –

- First, SVM will generate hyperplanes iteratively that segregates the classes in best way.
- Then, it will choose the hyperplane that separates the classes correctly.

CODE SNIPPETS:

1.) Loading of dataset in the program

In [2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
```

In [3]:
```
fertilizer_df=pd.read_csv("Fertilizer Prediction.csv")
fertilizer_df['Soil Type'].value_counts()
```

Out[3]:
```
Loamy     21
Sandy     20
Clayey    20
Red       19
Black     19
Name: Soil Type, dtype: int64
```

2.) Dividing data in training and test data and features selections also:

```
fertilizer_df.columns
feature_df=fertilizer_df[['Nitrogen', 'Potassium', 'Phosphorous']]
X=np.asarray(feature_df)
y=np.asarray(fertilizer_df['Fertilizer Name'])
```

```
from sklearn.model_selection import train_test_split
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,random_state=4)
```

3.) Importing and Training Model

In [7]:
```
from sklearn import svm
classifier=svm.SVC(kernel='linear',gamma='auto',C=2)
classifier.fit(X_train,y_train)
print("Enter soil nutritional components")
N=input("Nitrogen")
K=input("Potassium")
P=input("Phosphorous")
K_test=[[N,K,P]]
y_predict=classifier.predict(K_test)
print("Recommended Fertilizer:",y_predict)
y_predict1=classifier.predict(X_test)
```

OUTPUT

```
Enter soil nutritional components
Nitrogen12
Potassium13
Phosphorous14
Recommended Fertilizer: ['17-17-17']
```

```
Enter soil nutritional components
Nitrogen45
Potassium12
Phosphorous10
Recommended Fertilizer: ['Urea']
```

5. Result:

The connection was established for the option 1 and it works fine for it. But for option 2, We trained a Machine Learning Model and used that to predict the best fertilizers. Below is the prediction test result that came after when we checked our model for the test data with the code snippet too:

```
In [8]: from sklearn.metrics import classification_report
        print(classification_report(y_predict1,y_test))
```

```
              precision    recall  f1-score   support

    10-26-26       1.00      1.00      1.00         2
    14-35-14       1.00      1.00      1.00         3
       20-20       1.00      1.00      1.00         3
       28-28       1.00      1.00      1.00         2
         DAP       1.00      1.00      1.00         6
        Urea       1.00      1.00      1.00         4

    accuracy                           1.00        20
   macro avg       1.00      1.00      1.00        20
weighted avg       1.00      1.00      1.00        20
```

That was the result of the model when the training data was 80 percent and test data was 20 percent. On changing that ratio to 70 to 30 which is generally considered as the best proportion, the classification report changes by a little.

```
In [12]: from sklearn.metrics import classification_report
         print(classification_report(y_predict1,y_test))
```

```
              precision    recall  f1-score   support

    10-26-26       1.00      0.67      0.80         3
    14-35-14       0.80      1.00      0.89         4
    17-17-17       1.00      1.00      1.00         2
       20-20       1.00      1.00      1.00         3
       28-28       1.00      1.00      1.00         3
         DAP       1.00      1.00      1.00         9
        Urea       1.00      1.00      1.00         6

    accuracy                           0.97        30
   macro avg       0.97      0.95      0.96        30
weighted avg       0.97      0.97      0.97        30
```

Even after changing the ratio of distribution of training data and test data, the accuracy and precision we get in SVM is better than the other algorithms out there.

6. Future Challenges

The Significant Difficulties that can be experienced in future incorporate populace development, climatic changes, urbanization, expanding fulfillment of normal assets, herbicide obstruction, arising vermin and infections, changing food propensities, food misfortunes and wastage. In tropical districts, the test is to adjust to the methods of creation to the climatic fluctuation by attempting to lessen the dangers and weakness of the creation frameworks and additionally it is important to limit the negative effects of farming in the climate in any case there is a shot at expanding the likelihood of having antagonistic occasions.

The Base Misfortunes from the antagonistic occasions can be acquired exclusively by fostering the danger pointers which permits expectations of the effect of dangers that must be joined by a productive correspondence system with a huge degree. In this way, the inner preparation of the creation units should be vigorous to carry out the moderate strategies, various creation exercises and specialized reception, yet for this there ought to be rural administrations organizations.

Since the tropical climate is ideal for nonstop issues with nuisances, weeds and infections. To stay away from the cataclysms from these issues, the yield food creature framework should be lined up with using biodiverse frameworks.With headways in AI and different procedures and strategies, fabricating new applications is conceptualized. Yet, the test in the specialized field is itself that the result of the organization's advancement as one of the rising requests for ecological ingestion of residuals is gotten from the development that is item creation and the utilization. A portion of the significant ideas and the imaginative examination thoughts are inventoried beneath:

1. There is a need in creating complex thoughts and techniques dependent on conditions like science, industry, skill and experience.

2. The Presentation of frameworks that can endure different kinds of soils and ecological conditions. As the horticultural terrains are cruel and continue to change either dependent on the climatic conditions or seasons winning.

3. The Improvement of fundamental frameworks that are vigorous and economical to the progressions in the outer also as interior elements.

4. The Advancement of different sensor-based frameworks that are ended up being proficient for exceptionally raised regions that can for the most part have swaying surfaces and furthermore in many nations it is as yet a fantasy to supply capacity to such
regions.
5. The Utilization of over a wide span of time information for fostering the end taking in emotionally supportive networks for horticulture. Anyway by and large the planning of past cultivating engineering, the significant information is just taken into
count.

7. Conclusion:

The proposed site design is characterized by attractive and provides an easy navigation, limited option in term of brand, color and design. The purpose of this paper is to locate the customer's flexible, attractive and easy-to-use websites to provide them an option of choosing the best fertilizers on the basis of the crop type, soil type or crop nutritional requirement. The work can be extended further to add functionality like crop production yield, solving general queries of farmer, updating farmers with latest Government Schemes or telling them what is trending etc . Mobile applications can be built to help farmers by uploading image of farms. Crop diseases detection using image processing in which user get pesticides based on disease images. Implement Smart Irrigation System for farms to get higher yield.

8. Reference:

1.) Dr..G.Suresh, Dr.A.Senthil Kumar, Dr.S.Lekashri and Dr.R.Manikandan(2021) Efficient Crop Yield Recommendation System Using Machine Learning For Digital Farming .IJMA

2.) K. Dhivyaa, M. Malarvizhi, Dr.M. Malleswari(2021) Agriculture Marketing Recommendation System. IRJMETS

3.) https://www.researchgate.net/profile/Anantha-Reddy-Dasari/publication/350089918_Crop_recommendation_system_to_maximize_crop_yield_in_ramtek_region_using_machine_learning/links/60507400299bf1736746a448/Crop-recommendation-system-to-maximize-crop-yield-in-ramtek-region-using-machine-learning.pdf

4.) https://www.turcomat.org/index.php/turkbilmat/article/view/1951

5.) E. Manjula, and S. Djodiltachoumy (2017). Data Mining Technique to Analyze Soil Nutrients based on Hybrid Classification. IJARCS

6.) http://ijeast.com/papers/371-376,Tesma405,IJEAST.pdf

7.) https://www.proquest.com/openview/c0a5ba585e1e0d4d037079e468d6d979/1?pq-origsite=gscholar&cbl=2040552

8.) Veeresh Kadlimatti, S V Saboji, "Agriculture Price Prediction Using Data Mining"

9.) Devika B and Ananthi B, "Analysis of crop yield prediction using data mining technique to predict annual yield of major crops"

10.) Monali P, Santosh K, Vishwakarma, and Ashok V (2015). Analysis of Soil Behaviour and Prediction of Crop Yield using Data Mining Approach, ICCICN 2015.

11.) Mrs. N. Hemageetha, and Dr. G.M. Nasira (2016). Analysis of Soil Condition based on pH value using Classification Techniques, IOSR-JCE, Issue 6, (pp.50-54).

12.) E. Manjula, and S. Djodiltachoumy (2017). Data Mining Technique to Analyze Soil Nutrients based on Hybrid Classification. IJARCS.

13.) Rohit Kumar Rajak, Ankit Pawar, Mitalee Pendke, Pooja Shinde, Suresh Rathod, and Avinash Devare (2017). Crop Recommendation System to maximize Crop yield using Machine Learning. Issue 12 IRJET.

14.) S. Veenadhari, Dr. Bharat Misra, Dr. CD singh (2014). Machine Learning Approach for Forecasting Crop Yield based on Climatic Parameters. ICCCI-2014

15.) Sadia A, Abu Talha K, Mahrin Mahia, Wasit A, and Rashedur M.R.(2018). Analysis of Soil Properties and Climatic Data To Predict Crop Yields and Cluster Different Agricultural Regions of Bangladesh, IEEE ICIS 2018 (pp.80-85).

16.) Arun K, Navin K, and Vishal V (2018). Efficient Crop Yield Prediction using Machine Learning, (pp.3151-3159), IRJET.

17.) Subhadra M, Debahuti M, Gour, H. S.(2016). Applications of Machine Learning Techniques in Agricultural Crop Production: A Review Paper, Vol 9(38), DOI:10.17485/ijst/2016/v9i38/95032 IJST.

18.) Mrs. Prajakta P.B., Yogesh S. P, and Dinesh D.P. (2017). Improved Crop Yield Prediction using Neural Network, IJARIIE, (pg.3094-3101).

19.) Omkar B,Nilesh M,Shubham G,Chandan M, and D.S. Zingade (2017), Crop Prediction System using Machine Learning, Volume 4, Special Issue 5, IJAERD.

20.) Plant growth and soil moisture relationships, J.FBierhuizen, Unesco/NS/AZ/476, Madrid Symposium Paper N0.37, Paris 1 September 1959.
21.) Santhi, P., Mahalakshmi, G., Classification of magnetic resonance images using eight directions gray level co-occurrence matrix (8dglcm) based feature extraction, International Journal of Engineering and Advanced Technology, 2019, 8(4), pp. 839–846.
22.) Influence of Autumn Tilth on Soil Physical Conditions andTiming of Crop Establishment in Spring J. agric. Engng Res. (1984) 29, 265-213.
23.) Ramesh A, Vijay. S. Rajpurohit, "A survey on Data Mining Techniques for Crop Yield Prediction", "International Journal of Advance Research in Volume 2, Issue 9, September 2014".
24.) . Deepa, K., Kokila, M., Nandhini, A., Pavethra, A., Umadevi, M. "Rainfall prediction using CNN", International Journal of Advanced Science and Technology, 2020, 29(7 Special Issue), pp. 1623–1627. http://sersc.org/journals/index.php/IJAST/article/view/10849.