Crop Recommender System

using Machine Learning

Dr.S Sasi Kumar

Professor, *Department of Computer Science and Engineering,*

Saveetha Engineering College Chennai, Tamil Nadu, India.

[sasikumar@savvetha.ac.in](mailto:sasikumar@savvetha.ac.in)

K J PoornaKrishnaa

*Department of Computer Science and Engineering*

Saveetha Engineering College, Chennai, Tamil Nadu, India.

[poorna1803@gmail.com](mailto:poorna1803@gmail.com)

***Abstract -- Agriculture is an important factor in the Indian economy. A common problem faced by farmers in India is that they do not choose the right crop according to their soil requirements. As a result, their productivity drops significantly. This problem of farmers is solved by precision agriculture. Precision agriculture is a modern farming technique that uses research data on soil properties, soil types, crop data and recommends the right crop to farmers based on their location. parameters. This reduces wrong cropping and increases productivity. In this project, we are building an intelligent system to help farmers in India make informed decisions about which crop to grow based on the crop season, the geography of their farm and soil characteristics. In addition, the system provides the farmer with a crop forecast when he plants the recommended crop..***

***Keywords: Precision Agriculture, yield prediction, Dataset, rainfall, pH, NPK ratio.***

1. INTRODUCTION

A farmer's decision about which crop to grow is usually clouded by his intuition and other irrelevant factors such as immediate profit, ignorance of market demand, overestimation of the soil's potential to support a particular crop, etc. A very wrong decision by a farmer can seriously damage his family's finances.Perhaps this is one of the many reasons that contribute to the countless farmer suicides we hear about in the media every day. In a country like India, where agriculture and allied sectors account for about 20.4 percent of its Gross Value Added (GVA) [2], such a misjudgment would have a detrimental effect not only on the farmer's family but also on the economy as a whole. . region That is why we defined the farmer's dilemma as a very serious one, which crop to grow this season.. The need of the hour is to design a system that could

Provide Indian farmers with predictive knowledge and help them make informed decisions about which crops to grow. With this in mind, we propose a system, an intelligent system that takes into account environmental parameters.

(temperature, rainfall, geographical location in terms of state) and soilcharacteristics (pH value, soil type and nutrients concentration) beforerecommending the most suitable crop to the user.

1.1. Existing System

More and more scientists are beginning to understand this problem in Indian agriculture and are devoting more and more time and effort to mitigating the problem. Various jobs include using Regularized Greedy Forest to determine the appropriate processing order at a given timestamp. Another approach proposes a model that uses historical meteorological data as a training set. The model is trained to identify weather conditions that hinder apple production. It then effectively predicts the apple harvest based on monthly weather conditions. The use of several algorithms (e.g. Artificial Neural Network, K Nearest Neighbors and Regularized Greedy Forest) was demonstrated in [5] for crop selection based on yield prediction, which in turn is influenced by several parameters

.

* + 1. Drawbacks

One shortcoming we found in all these notable published works was that the authors of each paper focused on a single parameter (either weather or soil) to predict crop growth suitability. But in our opinion, both factors should be considered together to get the best and most accurate forecast. This is because a certain type of soil may be suitable for supporting one crop, but if the weather conditions in the area are not suitable for that type of plant, their yield will suffer.

1.2 Proposed System

To overcome the above shortcomings, we propose anintelligent crop recommendation system that considers all relevant parameters including temperature,precipitation, location and soil condition to predict crop suitability. This system is mainly related to the main task of the Agro-Consultant to provide crop recommendations to the algorithms of the farmers. We also provide analysis of the results of crops grown in different countries, which gives the user a simple and reliable overview of crop selection and planning..

* 1. Plan of Implementation

The steps involved in this system implementation are:-

1. Acquisition of Training Dataset:The accuracy of any

machine learning algorithm depends on the number of parameters and the correctness of thetraining dataset.For the system, we are using various datasets all downloadedfor government website and kaggle.

Datasets include:-

a) Data set of crop cost per hectare for main crops in each state

b) Data pre-processing: This step involves replacing zero and 0 values ​​of yield with -1 so that it does not affect the overall forecast. Additionally, we had to encode the dataset to feed it to the neural network.

c) Training model and threshold recommendation: After the pre-processing step, we used the dataset to train different machine learning models such as neural network and linear regression. to achieve the highest possible accuracy..

Problem Statement

The inability of farmers to use traditional and unscientific methods to select the best crop for their land is a major problem in a country where about 50 percent of the population is engaged in agriculture. Availability and availability of correct and up-to-date information prevents potential researchers from engaging in case studies in developing countries. We proposed a system that can solve this problem by providing predictive information about crop recommendations based on trained machine learning models. taking into account main environmental and economic indicators..

Objective of the Project

• To build a robust model to give correct and accurate prediction of crop sustainability in a given state for the particular soil type and climaticconditions.

• Provide recommendation of the best suitable crops in the area so that thefarmer does not incur any losses

• Provide profit analysis of various crops based on previous years data

1. LITERATURESURVEY
2. Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique Authors: Rakesh Kumar, M.P. Singh, Prabhat Kumar and J.P.Singh.

• 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.

1. AgroConsultant: Intelligent Crop Recommendation System Using MachineLearning Algorithms Authors: ZeelDoshi, SubhashNadkarni, RashiAgrawal, Prof. Neepa Shah.

• This paper, proposed and implemented an intelligent crop recommendation system, which can be easily used by farmers all over India. This systemwould assist the farmers in making an informed decision about which crop to grow depending on a variety of environmental and geographical factors. Wehave also implemented a secondary system, called Rainfall Predictor, which predicts the rainfall of the next 12 months.

1. Development of Yield Prediction System Based on Real-time Agriculturalmeteorological Information Haedong Lee \*, Aekyung Moon\* \* ETRI, 218Gajeong-ro, Yuseong-gu, 305-700, Korea

* This paper contains about the research and the building of an effectiveagricultural yield forecasting system based on real-time monthly weather. It isdifficult to predict the agricultural crop production because of the abnormalweather that happens every year and rapid regional climate change due toglobal warming. The development of agricultural yield forecasting systemthat leverages real-time weather information is urgently required. In thisresearch, we cover how to process the number of weather data(monthly, daily)and how to configure the prediction system. We establish a non-parametricstatistical model on the basis of 33 years of agricultural weather in- formation.According to the implemented model, we predict final production using themonthly weather information. This paper contains the results of thesimulation.

1. Analysis of Soil Behaviour and Prediction of Crop Yield using Data MiningApproach Monali Paul, Santosh K. Vishwakarma, Ashok Verma Computerscience and Engineering GGITS, Jabalpur

* This work presents a system, which uses data mining techniques in orderto predict the category of the analyzed soil datasets. The category, thuspredicted will indicate the yielding of crops. The problem of predicting thecrop yield is formalized as a classification rule, where Naive Bayes andK-Nearest Neighbor methods are used.

1. Crop Recommendation System for Precision Agriculture S.Pudumalar\*,E.Ramanujam\*, R.HarineRajashree, C.Kavya, T.Kiruthika, J.Nisha.

* This paper, proposes a recommendation system through an ensemblemodel with majority voting technique using Random tree, CHAID, K-NearestNeighbor and Naive Bayes as learners to recommend a crop for the sitespecific parameters with high accuracy and efficiency.

1. METHODOLOGY
2. Data Analysis

One of the first steps we take during implementation is data analysis. We did this to try to find possible relationships between different attributes of the dataset.Getting the training dataset: The accuracy of any machine learning algorithm depends on the number of parameters and the accuracy of the training dataset. . In this project, weanalyzed two datasets collected from the government website - https://data.gov.in/ and Kaggle - and carefully selected the parameters that would give the best results. Much of the work done in this field has considered environmental protection.

Some took yield as the main factor to predict ramet resistance, while, as in some work, only economic factors were considered. We have tried to combine environmental parameters such as rainfall, temperature, pH, soil nutrients, soil type, location and economic parameters such as production and yield to give the farmer an accurate and reliable recommendation on which crop is best suited for his land.

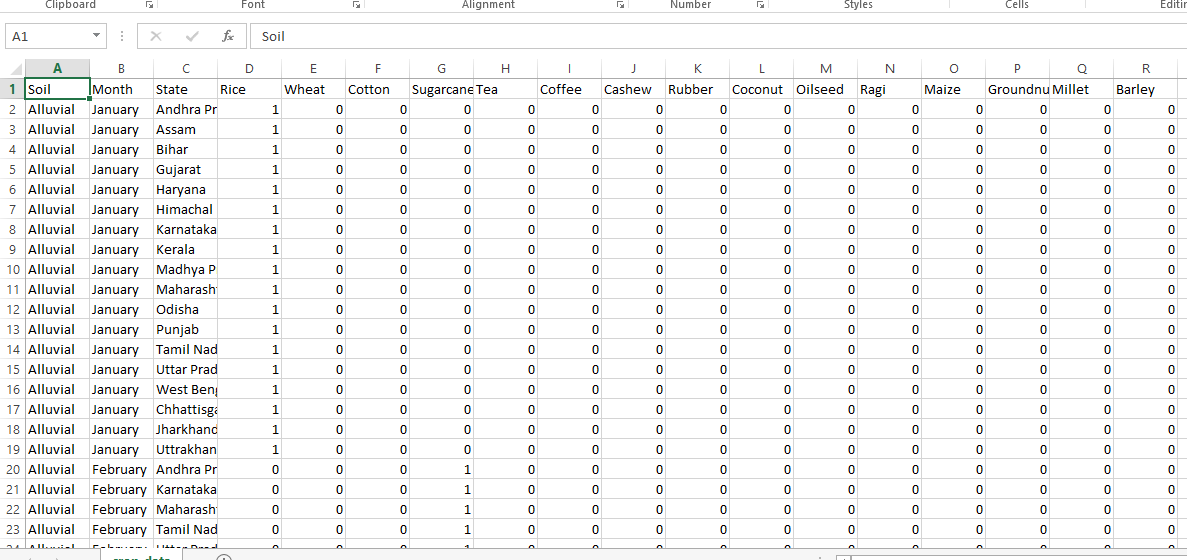


Fig.1. Encoded dataset

### Profit analysis was performed using cost of cultivation, market price, standard price and yield dataset. This was performed as a first step to know how much impact profit as a parameter can has on crop prediction.

1. Data Preprocessing

After data analysis and visualization, the next step is preprocessing. Data preprocessing is an important step because it helps to clean the data and make it suitable for use in machine learning algorithms. Most of the preprocessing focuses on removing abnormal or incorrect data and processing missing values.Missing data can be handled in two ways. The first option is to simply delete an entire row that contains a missing or invalid value. Although this method is easy to perform, it is best to use it only for large data sets. Using this method on small data sets can reduce the size of the data set too much, especially if there are many missing values. This can significantly affect the accuracy of the result. Since our dataset is relatively small, we do not use this method.The dataset we used had values ​​that were in string format, so we had to convert and encode them into integers to pass them as input to the neural network. We first converted the data into panda category data and then generated harvest and state codes, combined them and created separate datasets. To reduce the amount of data for the linear regression model, we filtered grains for required nutrients and nutrients in the soil. If the nutrient content of the soil fell below the amount needed by the crops, the crop was discarded, which allowed us to significantly reduce the training time.

# Machine Learning Models

Linear regression is a linear approach to modeling the relationship between a scalar response (or dependent variable) and one or more explanatory variables (or independent variables). Linear regression is used to find a linear relationship between a subject and one or more predictors. It fits a linear model with coefficients w = (w1, ... ,wp) to minimize the residual sum of squares between the observed objects in the data set and the objects predicted by the linear approximation. Linear regression is used to find a linear relationship between a subject and one or more predictors.Neural networks are a set of algorithms, loosely modeled after the human brain, designed to detect patterns.

They interpret sensory data using a form of machine perception, labeling or summarizing the raw input. The patterns they detect are numerical and contained in vectors into which all real data, whether images, audio, text or time series, must be translated..

1. TESTING

# Testing Methodologies:

The program consists of several algorithms, the accuracy of which is tested separately. We verify the correctness of the program as a whole and its performance.Unit testing: Unit tests focus on ensuring that the correct changes to the world state occur during event processing. The business logic of event processing functions should have unit tests, preferably with 100% code.

This ensures that there are no typos or logic errors in the motion logic. Different modules can be run separately from the command line and tested for correctness.The tester can pass different values ​​to check the returned response and check it against the values ​​given to it. Another option is to write a script and run all the tests with it and write the output to a log file and use it to check the results. We tested each algorithm individually and modified the preprocessing accordingly to increase accuracy.

# System Testing:

System testing is a level of software testing where complete and integrated software is tested. The purpose of this test is to assess whether the systems meet the defined requirements. System testing is the testing of complete and fully integrated software and White Box Testing. System testing falls under the category of software black box testing.

Different types of system testing:

• Usability testing - Usability testing mainly focuses on the ease of use of an application by users, the flexibility of using controls and the ability of the system to achieve its goals.

• Testing of load - Load testing is necessary to know that the software solution works under real loads.

• Regression testing - - Regression testing involves testing that is done to ensure that none of the changes made during the development process lead to new errors.

• Recovery testing - Recovery testing is performed to show that the software solution is reliable, trustworthy and can successfully recover from potential crashes.

• Migration testing - the migration testing is performed to ensure that software can be migrated from older system infrastructures to the present. system infrastructure without problems.

# Quality Assurance:

# Quality assurance, commonly known as QA testing, is defined as the activity of ensuring that an organization provides the best possible product or service to its customers. QA focuses on improving processes to deliver quality products to the customer. The organization must ensure that the processes are efficient and effective according to the quality standards defined for the software products.

# Functional Test:

Also known as functional completeness testing, functional testing involves coming up with potential missing features. As the chat room evolves into new application areas, for functional testing of

important chatbot components, Functional Testing evaluates use case scenarios and related business processes, such as the behavior of smart contracts.

IV. RESULT AND DISCUSSION

### For this project we used three popular algorithms: linear regression, logistic regression and neural network. All algorithms are based on supervised learning. Our entire system is divided into one module:. Crop recommender

Figure1.system architecture

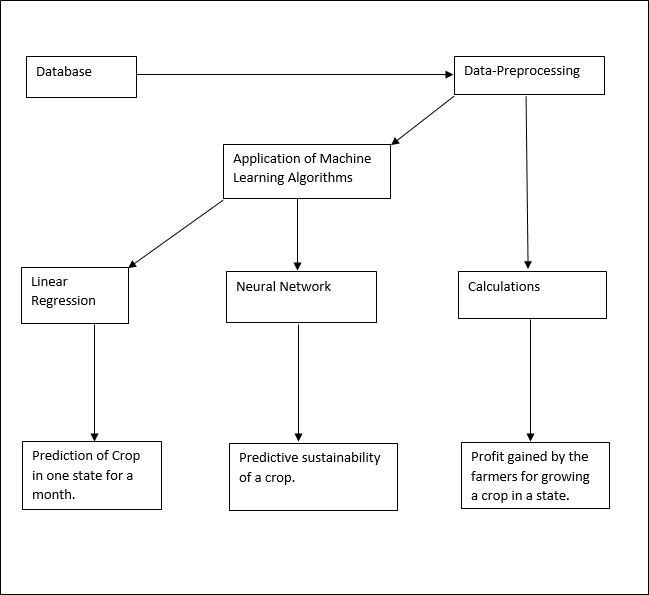


Figure 1 shows the system architecture of the project.

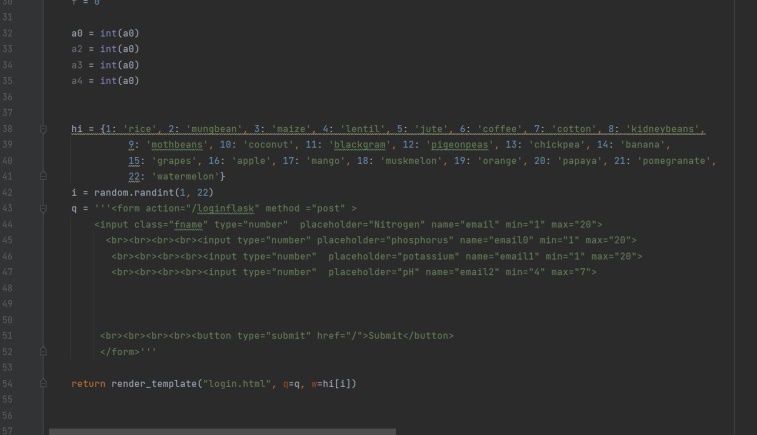
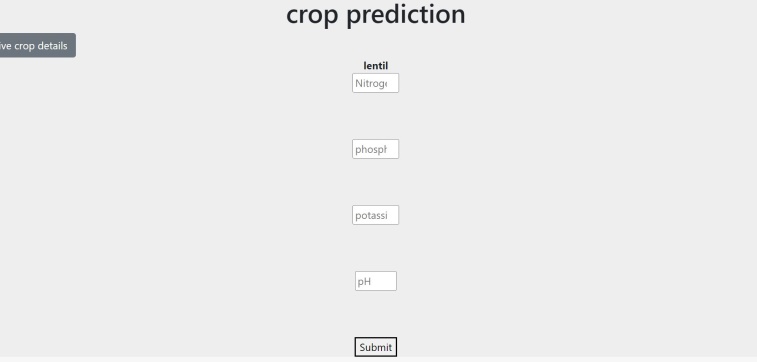


Figure 2 shows the partial program of implementation in the project.



This is the first output that obtained when we give certain values for the npk ratios with the pH value of the soil.

**v. CONCLUSION**

This system helps farmers choose the right crop by providing insights that are not observed by ordinary farmers, reducing the possibility of crop failure and increasing productivity. It also prevents taking losses. The system can be scaled online and used by millions of farmers across the country. We could get 89.88 percent accuracy from the neural network and 88.26 percent accuracy from the linear regression model. A further development is to integrate the crop recommendation system with another subsystem, a yield predictor, which would also provide the farmer with an estimate of production when he plants the recommended plants.

**References:**

### RakeshKumar , M.P. Singh, Prabhat Kumar and J.P. Singh, “Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique”, International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials, 2015

### Haedong Lee and Aekyung Moon, “Development of Yield Prediction System Based on Real-time Agricultural Meteorological Information”, 16th International Conference on Advanced Communication Technology, 2014

### T.R. Lekhaa, “Efficient Crop Yield and Pesticide Prediction for Improving Agricultural Economy using Data Mining Techniques”, International Journal of Modern Trends in Engineering and Science (IJMTES), 2016, Volume 03, Issue 10

### AnshalSavla, ParulDhawan, HimtanayaBhadada, Nivedita Is- rani, Alisha Mandholia, SanyaBhardwaj (2015), ‘Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture’, Innovations in Information, Embedded and Communication systems (ICIIECS).

### AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms. ZeelDoshi ,SubhashNadkarni , RashiAgrawal, Prof. Neepa Shah.

### Crop Recommendation System for Precision Agriculture S.Pudumalar\*, E.Ramanujam\*, R.HarineRajashree, C.Kavya, T.Kiruthika, J.Nisha.

### <https://data.gov.in/g>

### Kaggle<https://www.kaggle.com/notebook>

1. TESTING

# Testing Methodologies: The program comprises of several algorithms which are tested individually for the accuracy. We check for the correctness of the program as a whole and how it performs.

# Unit Testing:

# Unit tests focus on ensuring that the correct changes to the world- state take place when a transaction is processed. The business logic in transaction processor functions should have unit tests, ideally with 100 percent code coverage. This will ensure that you do not have typos or logic errors in the business logic. The various modules can be individually run from a command line and tested for correctness. The tester can pass various values, to check the answer returned and verify it with the values given to him/her. The other work around is to write a script, and run all the tests using it and write the output to a log file and using that to verify the results. We tested each of the algorithms individually and made changes in preprocessing accordingly to increase the accuracy.

# System Testing:

# System Testing is a level of software testing where complete and integrated software is tested. The purpose of this test is to evaluate the systems compliance with the specified requirements. System Testing is the testing of a complete and fully integrated software product and White Box Testing. System test falls under the black box testing category of software testing. Different Types of System Testing:

### Usability Testing - Usability Testing mainly focuses on the users ease to use the application, flexibility in handling controls and ability of the system to meet its objectives.

### Load Testing - Load Testing is necessary to know that a software solution will perform under real-life loads.

### Regression Testing- - Regression Testing involves testing done to make sure none of the changes made over the course of the development process have caused new bugs.

### Recovery Testing - Recovery testing is done to demonstrate a soft- ware solution is reliable, trustworthy and can successfully recoup from possible crashes.

### Migration Testing - Migration testing is done to ensure that the software can be moved from older system infrastructures to current system infrastructures without any issues.

# Quality Assurance:

# Quality Assurance is popularly known as QA Testing, is defined as an activity to ensure that an organization is providing the best possible product or service to customers. QA focuses on improving the processes to deliver Quality Products to the customer. An organization has to ensure, that processes are efficient and effective as per the quality standards defined for software products.

# Functional Test:

### Functional Testing is also known as functional completeness testing, Functional Testing involves trying to think of any possible missing functions. As chat-bot evolves into new application areas, functional testing of essential chatbot components, Functional testing evaluates use-case scenarios and related business processes, such as the behavior of smart contracts.

1. RESULT AND DISCUSSION

For the purposes of this project we have used three popular algorithms: Linear regression, Logistic regression and neural network. All the algorithms are based on supervised learning. Our overall system is divided into 1 module:

### Crop recommender

Figure1.system architecture

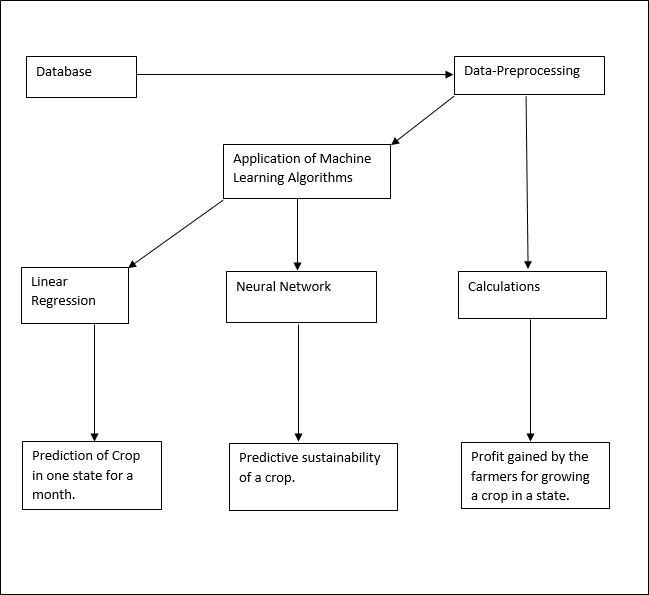
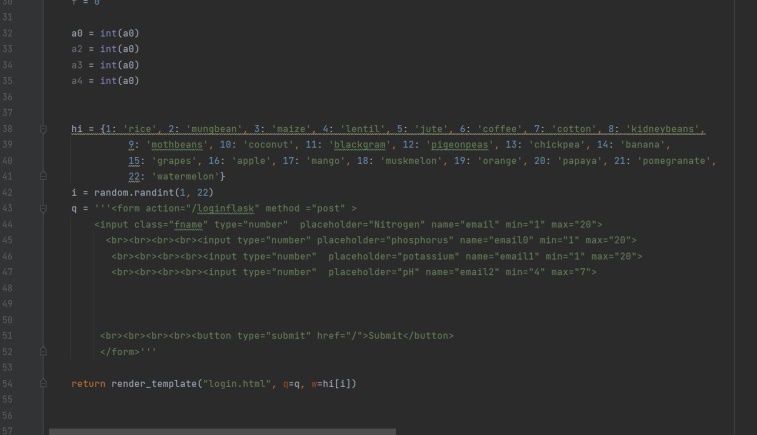
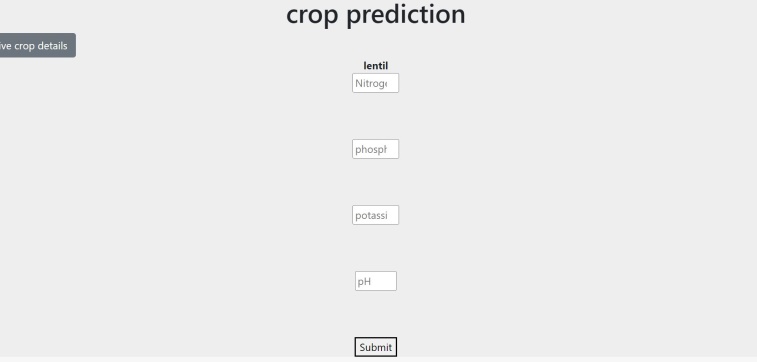


Figure 1 shows the system architecture of the project.

Figure 2 shows the partial program of implementation in the project.





This is the first output that obtained when we give certain values for the npk ratios with the pH value of the soil.

**v. CONCLUSION**

This system helps the farmer to choose the right crop by providing insights that ordinary farmers don’t keep track of thereby decreasing the chances of crop failure and increasing productivity. It also prevents them from incurring losses. The system can be extended to the web and can be accessed by millions of farmers across the country. We could achieve an accuracy of 89.88 percent from the neural network and an accuracy of 88.26 percent from the linear regression model**.** Further development is to integrate the crop recommendation system with another subsystem, yield predictor that would also provide the farmer an estimate of production if he plants the recommended crops.

**References:**

### RakeshKumar , M.P. Singh, Prabhat Kumar and J.P. Singh, “Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique”, International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials, 2015

### Haedong Lee and Aekyung Moon, “Development of Yield Prediction System Based on Real-time Agricultural Meteorological Information”, 16th International Conference on Advanced Communication Technology, 2014

### T.R. Lekhaa, “Efficient Crop Yield and Pesticide Prediction for Improving Agricultural Economy using Data Mining Techniques”, International Journal of Modern Trends in Engineering and Science (IJMTES), 2016, Volume 03, Issue 10

### AnshalSavla, ParulDhawan, HimtanayaBhadada, Nivedita Is- rani, Alisha Mandholia, SanyaBhardwaj (2015), ‘Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture’, Innovations in Information, Embedded and Communication systems (ICIIECS).

### AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms. ZeelDoshi ,SubhashNadkarni , RashiAgrawal, Prof. Neepa Shah.

### Crop Recommendation System for Precision Agriculture S.Pudumalar\*, E.Ramanujam\*, R.HarineRajashree, C.Kavya, T.Kiruthika, J.Nisha.

### <https://data.gov.in/g>

### Kaggle<https://www.kaggle.com/notebook>