Crop Recommendation & Disease Detection using Machine Learning

Sameer Panat[1] Sarthak Bhaiji[2] Siddhant Tripathi[3] Vandana Kate[4] Kavita Namdev[5]
[1,2,3]Student [4]Associate Professor [5]Senior Assistant Professor
[1,2,3,4,5]Department of Computer Science and Engineering
[1,2,3,4,5]Acropolis Institute of Technology and Research, Indore, MP, India

Abstract— In this ever-transmuting world of technology to make farmers go hand in hand with the vicissitude, we have developed a product that consummates the gap between farmer and incipient technological advancement. Our product is a one-click solution for the farmer predicated on machine learning. Suggesting farmer most felicitous crop and fertilizer according to its soil conditions. The project is a web application, based on machine learning algorithms like CNN, Deep neural nets, etc, analyzing different key nutrients present in the soil and utilize it to recommend congruous crops predicated on the soil contents. This product removes the dependency on soil health cards and makes the process more fluid and quick. This product can also be used to detect disease in the plant just by its leaf image making it a complete solution for agricultural practices.

Keywords: Machine Learning, CNN, MLP, Crop Recommendation, Disease Detection

I. INTRODUCTION

The agriculture sector is one of the most important sectors in many countries. The agriculture sector provides food, raw material, and employment. The agriculture sector faces many problems such as irregular rainfall, floods, drought, climate change, etc. To overcome these problems technological solution is needed which can help the farmers. The productivity of farming is not only depending on natural resources, but it also depends on input provided to the system such as type of soil, type of fertilizer, and crops. These inputs can increase or decrease the productivity of any crop. old age crop selection is not precise and does not have any analysis details. To surmount such quandaries data mining and machine learning techniques can be utilized. In the agriculture sector data set for a long time duration is available. Data mining is used to analyze the dataset and extract information from it. Machine Learning algorithms can be used to predict the proper crop so that it will lead to less loss in inputs and increases the overall profit.

We live in a country whose 60% of the population is directly involved in agriculture the potentiating them would potentiate the nation's magnification. Cultivating a crop is not everyone's cup of tea, as it is a very tedious task and requires constant attention towards crop, due to incongruous utilization of pesticides and other prominent reasons crop suffer sundry diseases that need a timely and precise diagnosis of diseases. It is crucial to obviate nonessential waste of financial and other resources, thus achieving more salubrious engenderment in this transmuting environment. Congruous and timely disease identification including early aversion has never been more paramount. There are several ways to detect plant pathologies. Some diseases do not have any visible symptoms, or the effect becomes conspicuous too tardy to act, and in those situations, a sophisticated analysis is obligatory. However, most diseases engender manifestation within the color spectrum, therefore the unclad ocular perceiver examination of a trained professional is that the prime technique adopted in practice for disease detection. so as to realize precise diagnostics, a plant pathologist should possess good optical discernment skills so that one can identify characteristic symptoms. Variations in symptoms betokened by diseased plants may lead to an incongruous diagnosis since tyro gardeners and hobbyists could have more difficulties determining it than a professional plant pathologist.

A Machine Learning System is required to avail identify plant diseases by the plant's appearance and visual symptoms could be of great avail to abecedarians in the farming process. Advances in computer vision present an opportunity to expand and enhance the practice of precise plant bulwark and elongate the market of computer vision applications in the field of precision agriculture. Exploiting prevalent digital image processing techniques such as color analysis and thresholding was utilized with the aim of detection and relegation of plant diseases. According to disease, a remedy is suggested to remedy that disease.

II. LITERATURE SURVEY

The framework in paper [1] recommended by creators S. Pudumalar and related co-creators utilizes a gathering system called Majority Voting Technique which consolidates the intensity of numerous models to accomplish more prominent forecasts precision. The strategies utilized are Random Trees, KNN, CHAID and Naïve Bayes for gathering so that regardless of whether one strategy predicts erroneously, alternate models are probably going to make the right forecasts and since the greater part casting a ballot system is utilized, the last expectation is the right one. On the off chance that rules are the fundamental segments that are utilized in the expectation procedure. The exactness got is 88% utilizing the gathering model.

Paper [2] is an audit paper to concentrate on different calculations and their exactness in the rural field proposed by Yogesh Gandge and Sandhya. It was seen that Multiple Linear Regression gave a precision of 90-95% for rice yield. Choice tree utilizing ID3 calculation was considered for soybean edit and the proposals were created. The third calculation was SVM which was utilized on every one of the harvests and the precision was great with computationally fewer prerequisites. The neural system was utilized on corn information to accomplish 95% of precision. Different calculations were additionally utilized which are KNN, C4.5, K-implies, J48, LAD Tree and Naïve Bayes. The end was that still enhancement is required for the calculations to accomplish better precision.

Being used as Data Mining in Crop Yield Prediction [3], paper [3], the dataset utilized was gathered from Kaggle.com The creator has broken down the information utilizing the WEKA apparatus for calculations

which are LWL, J48, LAD Tree and IBK. The exactness was estimated utilizing explicitness, affectability, precision, RMSE and mean outright blunder. For every classifier, the perplexity network was utilized to get the effectively distinguished occurrences. The perception was that better precision can be acquired if pruning is utilized.

Paper [4] displayed by Rakesh Kumar, M.P. Singh, Prabhat Kumar, and J.P. Singh proposed utilization of seven machine learning methods, for example, ANN, SVM, KNN, Decision Tree, Random Forest, GBDT and Regularized Gradient Forest for crop choice. The framework is intended to recover every one of the yields sowed and time of developing at a specific time. The yield rate of each harvest is acquired and the crops giving higher yields are chosen. The framework additionally proposes an arrangement of crops to be planted to get higher yields.

Prof. Rakesh Shirsathand other co-creator in paper [5] proposed a framework that causes the clients to settle on choices for the crop to be planted. The framework utilized is a membership-based framework which would have customized data of each rancher enlisted. The framework incorporates a module that keeps up the data of the past harvests planted gathered from different sources and demonstrates a coordinating crop that can be planted. The entire procedure is finished with the assistance of fake impartial systems. Toward the end, a criticism framework is given with the goal that the designer can make changes required if the rancher discovers some trouble while utilizing the framework.

Enormous Data Analysis Technology Application in Agricultural Intelligence Decision System paper creators Jinchuan Zhao and Jian-Xin Guo in paper [6] considers the information database as large information and derivations from the information is drawn. It considers different modules like clients, information build, area master, man-machine interface, deduction motor, and learning base. The learning procurement framework gets information for the choice framework and sets up a powerful learning base to take care of the issue.

III. RESEARCH METHODOLOGY

Our product is a recommender system that recommends the congruous crop and detects plant disease to the farmers. Rudimentary, the project divides into multiple modules.

A. About Dataset:

Our Dataset for the Crop Recommender System comprises five most prominent factors of soil for production i.e Nitrogen, Phosphorus, Potassium, Organic Carbon, Ph Value. Nitrogen is responsible for the development of proteins in the plant, phosphorus is responsible for the plant's ability to store energy, potassium takes care of plant's resistance to disease, Organic Carbon takes care of plant's stability and aeration capacity. All these features vary according to different crops and different region's soil formation.

Data is collected from the Government Website for Soil Health Card.

Dataset of Plant disease is collected from Github which comprise of image pixels as a feature set to train model in detecting disease form image sample.

B. OCR Module (Deep Learning):

The first module is just a simple User Interface. This User interface is engendered with the avail of python flask. Through this Utilizer Interface, farmers upload their Soil Health Card image and our OCR module works to extract data from the soil health card like the values of Nitrogen, Phosphorous, Potassium, Carbon Oxide Percentage, pH and store them as a test feature set for our model. Data is extracted using Google Tesseract-OCR Engine which uses Convolutional Neural Network to extract text, a number from the image.

C. Crop Recommender Module (Machine Learning):-

In this Module, we utilized the soil health card dataset to train our model on different features like nitrogen, phosphorous, potassium, carbon oxide, pH value, and the crop name as a label. Here Rudimentary we have used Neural Network to prognosticate the all best felicitous crops for farmers. In relegation, we find the probability of each class and show those crops whose probability is more preponderant than the mean of the probability of all classes and show them as output. The network comprises 2 hidden layers where Relu (Rectified Linear Unit) function is used as the activation function for hidden layers neuron and at output layer Sigmoid function is used to get a normal probability distribution of crops.

D. Plant Disease Detector:

In this module, we have trained our model using a plant disease dataset which comprises image pixel set as features to evaluate. The model consists of 6 convolutional layers where Inverse Dropout is used to Regularize the model at the output layer Sigmoid function is used to get a normal probabilistic distribution to identify the disease from a list of disease. Through this not only we detect the disease but also recommend an organic solution to the problem.

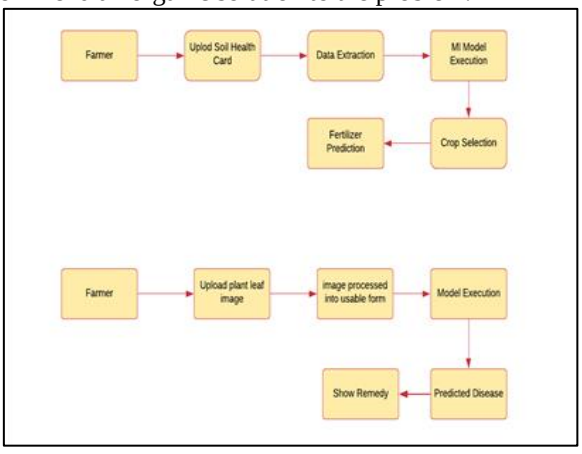

Fig. 3.1: Data Flow of Project

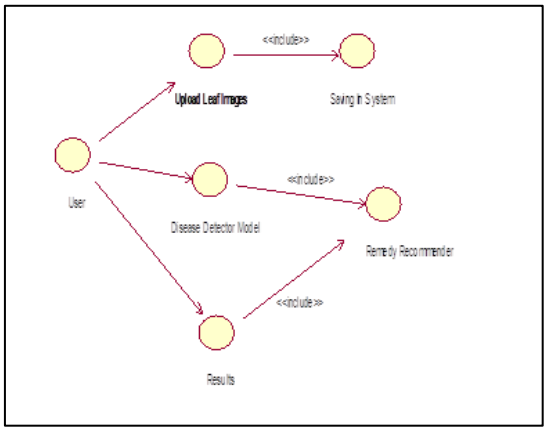

Fig. 3.2: Use case of Plant Disease Detection

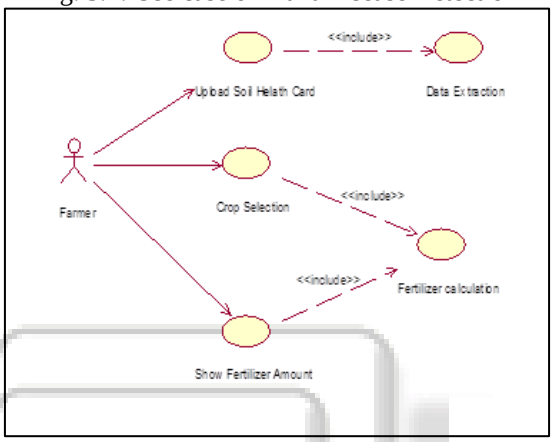

Fig. 3.3: Use case of Crop Recommender module

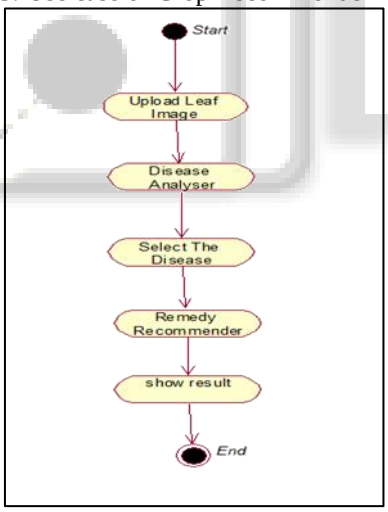

Fig. 3.4: Activity diagram

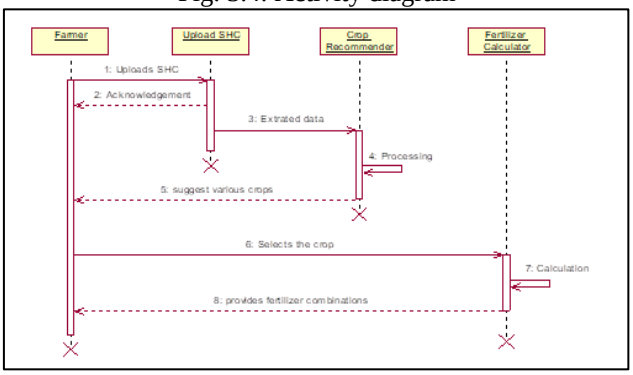

Fig. 3.5: Sequence Diagram

IV. RESULT

The prediction accuracy of the model accounts for 88%. The rules induced by the Multilayer perceptron model and Deep Convolutional network return a list of labels denoting the most courageous crop or disease in the plant from a normal probabilistic distribution of label classes.

The rules generated from the model are used to develop a RECOMMENDER SYSTEM. This is accomplished by creating a GUI. The GUI is deployed as a web-application. The model which is trained with the training data set is tested with inputs from the user. The scripting done will respond to any test case suggesting a crop, detecting a disease or recommending an optimum fertilizer combination.

Fig. 4.1: Home page of the application

Fig. 4.2: Option available to use.

Fig. 4.3: Crop output of the recommender module

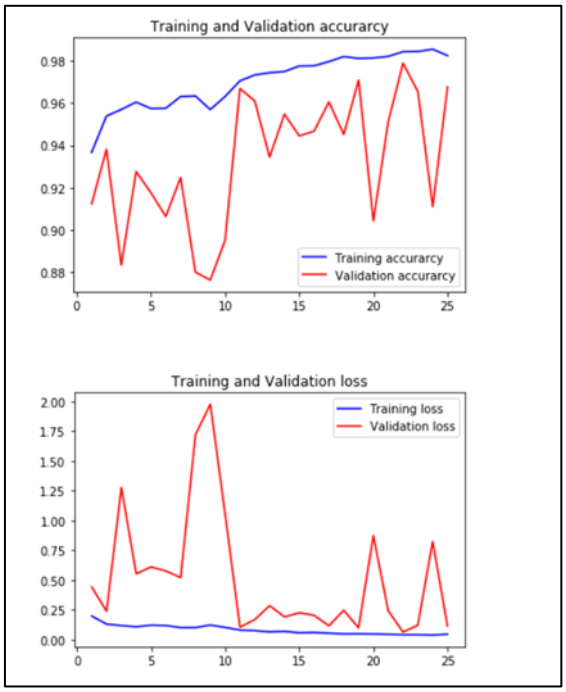

Fig. 4.4: learning curves of the model.

V. CONCLUSION

In today's world, people are magnetized to many fancy technologies which leave them awestruck and make them ruminate about the science behind it. India is an agricultural driven economy. In the prosperity of the farmers, prospers the nation. Thus our work would avail farmers in sowing the right seed predicated on soil requisites to increase productivity and acquire profit out of such a technique. Thus the farmers can plant the right crop incrementing his yield and withal incrementing the overall productivity of the nation.

A. Future Enhancement

- It can be implemented in all the cities.
- Multilingual Communication.
- Increase scope to all crops
- Remove dependency on labs
- Provide Farmer a genuine-time market rate of crops.

REFERENCES

[1] 2016 IEEE Eighth International Conference on Advanced Computing (ICoAC), "Crop Recommendation System for Precision Agriculture",
[2] 2017 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), "A Study on Various Data Mining Techniques for Crop Yield Prediction", Yogesh Gandge, Sandhya.
[3] 2017 IEEE Region 10 Humanitarian Technology Conference, "RSF: A Recommendation System for Farmers", Miftahul Jannat Mokarrama; Mohammad Shamsul Arefin.
[4] 2017 International Conference on I2C2, "Agriculture decision support system using data mining", Prof. Rakesh Shirsath; Neha Khadke; Divya More.
[5] 2018 the 3rd IEEE International Conference on Cloud Computing and Big Data Analysis, "Big Data Analysis Technology Application in Agricultural Intelligence Decision System", Ji-chun Zhao; Jian-xin Guo. Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018), "Use of Data Mining in Crop Yield Prediction", Shruti Mishra, Priyanka Paygude, Snehal Chaudhary, SonaliIdate.
[6] IEEE Sponsored 9th International Conference on Intelligent Systems and Control (ISCO) 2015, "XCYPF: A Flexible and Extensible Framework for Agricultural Crop Yield Prediction", Aakunuri Manjula, Dr.G .Narsimha.

