IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Healthy Harvesting Using KNN Algorithm

[1]G.Tejo Lakshmi, [2]Pavan Lokesh Tai, [3]T.Pavan Kalyan, [4]T.Jaswanth Jabez, [5]S.Yagna Koushik

[1]Assistant Professor, Department of CSE, KKR & KSR Institute of Technology and Sciences

[2,3,4,5] B.Tech Students, Department of CSE, KKR & KSR Institute of Technology and Sciences.

ABSTRACT: Agriculture is important and foremost factor which is important for survival. Machine Learning could be a crucial perspective for acquiring real-world and operative solution for crop yield issue. Machine Learning is a conspicuous decision support tool for crop yield prediction. this paper focusses mainly on predicting the yield of the crop by using by applying various machine learning algorithms. The farmer provides the attributes of the soil as input to the application. India being agriculturally based dependent country the economic status of the country is completely and partially dependent on this Agricultural yield that is affected by the organic, economic and seasonal causes. Estimation of agricultural output is big challenging task for this country as of the population status talking in consideration. The main objective is to gather data that can be stored and analyzed for forecasting the crop yield. This helps farmers to choose the best suitable crop. Also, this project aims at bringing an enhancement in the field of agriculture by achieving better results in predicting crop yields. With the help of machine learning techniques with proper optimizations. The statistical models built to provide accurate and precise decision. The output of this project would help farmers pick up most suitable crops to be grown depending on the factors like season and area available with least possible chances of loses.

Keywords: Agriculture, Crop yield, soil.

I. INTRODUCTION

Anthropogenic climate change will affect the agricultural sector more directly than many others because of its direct dependence on weather (Porter et al 2013). The nature and magnitude of these impacts depends both on the evolution of the climate system, as well as the relationship between crop yields and weather. This paper focuses on the latter—yield prediction from weather. Accurate models mapping weather to crop yields are important not only for projecting impacts to agriculture, but also for projecting the impact of climate change on linked economic and environmental outcomes, and in turn for mitigation and adaptation policy. A substantial portion of the work of modeling yield for the purpose of climate change impact assessment relies on deterministic, biophysical crop models (e.g. Rosenzweig et al 2013). These models are based on detailed representations of plant physiology and remain important, particularly for assessing response mechanisms and adaptation options (Ciscar et al 2018). However, they are generally outperformed by statistical models in prediction over larger spatial scales (Lobell and Burke 2010, Lobell and Asseng 2017). In particular, a large literature following Schlenker and Roberts (2009) has used statistical models to demonstrate a strong linkage between extreme heat and poor crop performance. These approaches have relied on classical econometric methods. Recent work has sought to fuse crop models with statistical models, variously by including crop model output within statistical models (Roberts et al 2017), and by using insights from crop models in the parameterization of statistical models (Roberts et al 2012, Urban et al 2015).

In parallel, machine learning (ML) techniques have advanced considerably over the past several decades. ML is philosophically distinct from much of classical statistics, largely because its goals are different—it is largely focused on prediction of outcomes, as opposed to inference into the nature of the mechanistic processes generating those outcomes. (We focus on supervised ML—used for prediction—rather than SNN and OLS projections greater than 16 percentage points in some of the most severe projections. Also, while OLS projections suggest yield increases in some northerly areas in some of the less severe scenarios, this projection is largely absent in the SNN projections. Finally, confidence intervals for mean projections are much smaller for the SNN than for the equivalent OLS regression, implying greater precision for any given weather scenario.

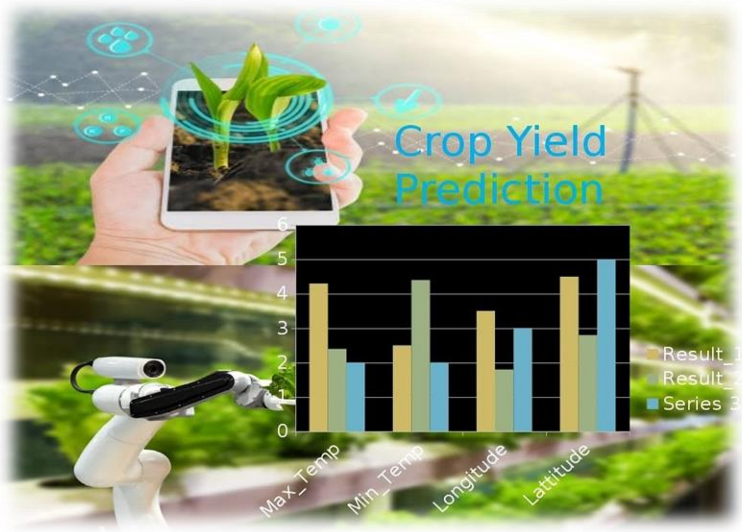

Fig1: crop yield prediction

II. LITERATURE REVIEW

In [1] Lack of automatic weed detection tools has hampered the adoption of site-specific weed control in cereals. An initial object-oriented algorithm for the automatic detection of broad-leaved weeds in cereals developed by SINTEF ICT (Oslo, Norway) was evaluated. The algorithm (‘‘Weed Finder’’) estimates total density and cover of broad-leaved weed seedlings in cereal fields from near-ground red–green–blue images. The ability of ‘‘Weed Finder’’ to predict ‘spray’/ ‘no spray’ decisions according to a previously suggested spray decision model for spring cereals was tested with images from two wheat fields sown with the normal row spacing of the region, 0.125 m. Applying the decision model as a simple look-up table, ‘‘Weed Finder’’ gave correct spray decisions in 65–85% of the test images. With discriminant analysis, corresponding mean rates were 84–90%. Future versions of ‘‘Weed Finder’’ must be more accurate and accommodate weed species recognition

In [2] this review, we present a comprehensive and critical survey on image-based plant segmentation techniques. In this context, ‘‘segmentation” refers to the process of classifying an image into plant and nonplan pixels. Good performance in this process is crucial for further analysis of the plant such as plant classification (i.e., identifying the plant as either crop or weed), and effective action based on this analysis, e.g., precision application of herbicides in smart agriculture applications. The survey briefly discusses pre-processing of images, before focusing on segmentation. The segmentation stage involves the segmentation of plant against the background (identifying plant from a background of soil and other residues). Three primary plant extraction algorithms, namely, (i) color index-based segmentation, (ii) threshold-based segmentation, (iii) learning-based segmentation are discussed. Based on its prevalence in the literature, this review focuses in particular on color index-based approaches. Therefore, a detailed discussion of the segmentation performance of color index-based approaches is presented, based on studies from the literature conducted in the recent past, particularly from 2008 to 2015. Finally, we identify the challenges and some opportunities for future developments in this space. In [3] Vegetables are a substantial part of our lives and possess great commercial and nutritional value Weeds not only decrease vegetable yield but also reduce their quality. Non-chemical weed control is important both for the organic production of vegetables and achieving ecologically sustainable weed management. Estimates have shown that the yield of vegetables may be decreased by 45%–95% in the case of weed–vegetable competition. Non-chemical weed control in vegetables is desired for several reasons. For example, there are greater chances of contamination of vegetables by herbicide residue compared to cereals or pulse crops.

[4] The injection of pesticide directly into spray nozzles results in concentration variations, both periodic discontinuities and pulses, caused by the use of injection pumps and solenoid valves. Such variations can lead to inaccurate application of pesticide. To avoid using long pipelines, long lag times and inaccurate application, efficient mixers are required to shorten the distance between the injection point and the spray nozzle. To analyze concentration variations and periodic pulses, a set of methods based on image processing were proposed and were tested using three inline mixers, a jet mixer (A), a layered mixer (B) which was specially designed, and simplified layered-jet mixer (C) which was developed as combination of mixer A and some simplified structures of mixer B. Results showed the characteristics of concentration variability for three mixers were generally different. [5]. A real-time intelligent robotic weed control system was developed for selective herbicide application to in-row weeds using machine vision and precision chemical application. The robotic vision system took 0.34s to process one image, representing a 11.43 cm by 10.16 cm region of seedling containing 10 plant objects, allowing the prototype robotic weed control system to travel at a continuous rate of 1.20 kmrh. The overall performance of the robotic system in a commercial processing tomato field and in simulated trials is discussed.

[6] Machine learning has emerged with big data technologies and high-performance computing to create new opportunities for data intensive science in the multi-disciplinary Agri-technologies domain. In this paper, we present a comprehensive review of research dedicated to applications of machine learning in agricultural production systems. The works analyzed were categorized in (a) crop management, including applications on yield prediction, disease detection, weed detection crop quality, and species recognition; (b) livestock management, including applications on animal welfare and livestock production; (c) water management; and (d) soil management. The filtering and classification of the presented articles demonstrate how agriculture will benefit from machine learning technologies. By applying machine learning to sensor data, farm management systems are evolving into real time artificial intelligence enabled programs that provide rich recommendations and insights for farmer decision support and action. [7] Weeds are among the major factors that could harm crop yield. With the advances in electronic and information technologies, machine vision combined with image processing techniques has become a promising tool for precise real-time weed and crop detection in the field, providing valuable sensing information for site-specific weed management. This review summarized the advances of weed detection using ground-based machine vision and image processing techniques. Concretely, the four procedures, i.e., pre-processing, segmentation, feature extraction and classification, for weed detection were presented in detail. To separate vegetation from background, different color indices and classification approaches like color index-based, threshold-based and learning-based ones, were developed. The difficulty of weed detection lies in discriminating between crops and weeds that often have similar properties. Generally, four categories of features, i.e., biological morphology, spectral features, visual textures and spatial contexts, were used for the task, which were discussed in this review. Application of conventional machine learning-based and recently developed deep learning-based approaches for weed detection were also presented. Finally, challenges and solutions provided by researchers for weed detection in the field, including occlusion and overlap of leaves, varying lighting conditions and different growth stages, were discussed.

In [8] most agricultural systems, one of the major concerns is to reduce the growth of weeds. In most cases, removal of the weed population in agricultural fields involves the application of chemical herbicides, which has had successes in increasing both crop productivity and quality. However, concerns regarding the environmental and economic impacts of excessive herbicide applications have prompted increasing interests in seeking alternative weed control approaches. An automated machine vision system that can distinguish crops and weeds in digital images can be a potentially cost-effective alternative to reduce the excessive use of herbicides. In other words, instead of applying herbicides uniformly on the field, a realtime system can be used by identifying and spraying only the weeds. This paper investigates the use of a machine-learning algorithm called support vector machine (SVM) for the effective classification of crops and weeds in digital images. Our objective is to evaluate if a satisfactory classification rate can be obtained when SVM is used as the classification model in an automated weed control system. In our experiments, a total of fourteen features that characterize crops and weeds in images were tested to find the optimal combination of features that provides the highest classification rate. Analysis of the results reveals that SVM achieves above 97% accuracy over a set of 224 test images. Importantly, there is no misclassification of crops as weeds and vice versa. In [9] An important objective in weed management is the discrimination between grasses (monocots) and broad-leaved weeds (dicots), because these two weed groups can be appropriately controlled by specific herbicides. In fact, efficiency is higher if selective treatment is performed for each type of infestation instead of using a broadcast herbicide on the whole surface. This work proposes a strategy where weeds are characterized by a set of shape descriptors (the seven Hu moments and six geometric shape descriptors). Weeds appear in outdoor field images which display real situations obtained from a RGB camera. Thus, images present a mixture of both weed species under varying conditions of lighting. In the presented approach, four decision-making methods were adapted to use the best shape descriptors as attributes and a choice was taken. This proposal establishes a novel methodology with a high success rate in weed species discrimination.

[10] Crop/weed recognition is a crucial step for selective herbicide application. A machine vision-based sensing system was developed to detect intra-row weeds when crops were at their early growth stages. The proposed methods used color feature to extract vegetation from the background, whilst height and plant spacing information analysis techniques were applied to discriminate between crops and weeds. Firstly, the identification of the weeds that were lower than crops was done by a height-based segmentation method using a stereo vision system. During the stereo matching process, correspondence search was performed on edged stereo images and disparity calculation was applied only to the edge pixels. This strategy could largely reduce the correspondence search range, thereby enhanced the weed recognition speed and accuracy. Afterwards, the higher weeds were distinguished from the crops by utilizing plant spacing characters. The histogram of plant pixels and their peak position were calculated from each pixel row of the segmented disparity image. Then plant centers were located and each weed region was further extracted based on the interplant distance in a row. [11] Leukemia is a fatal disease of white blood cells which affects the blood and bone marrow in human body. We deployed deep convolutional neural network for automated detection of acute lymphoblastic leukemia and classification of its subtypes into 4 classes, that is, L1, L2, L3, and Normal which were mostly neglected in previous literature. In contrary to the training from scratch, we deployed pretrained AlexNet which was fine-tuned on our data set. Last layers of the pretrained network were replaced with new layers which can classify the input images into 4 classes. To reduce overtraining, data augmentation technique was used. We also compared the data sets with different color models to check the performance over different color images. For acute lymphoblastic leukemia detection, we achieved a sensitivity of 100%, specificity of 98.11%, and accuracy of 99.50%; and for acute lymphoblastic leukemia subtype classification the sensitivity was 96.74%, specificity was 99.03%, and accuracy was 96.06%. Unlike the standard methods, our proposed method was able to achieve high accuracy without any need of microscopic image segmentation.

[12] Leukemia is the abnormal and uncontrolled development of the white blood cells, known as leukocytes, in the blood. The manual methods used for counting the blast cells have some demerits, and so automatic method must be employed. This paper proposes the Salp Swarm integrated Dolphin Echolocation-based Support Vector Neural Network (SSDE-SVNN) classifier to detect leukemia in its early stages. The pre-processed blood smear image is subjected to segmentation with the use of LUV transformation and Adaptive thresholding. [13] Despite the high prevalence of leukemia, there is a shortage of flow cytometry equipment, and the methods available at laboratory diagnostic centers are timeconsuming. Motivated by the capabilities of machine learning (machine learning (ML)) in disease diagnosis, the present systhematic review was conducted to review the studies aiming to discover and classify leukemia by using machine learning. Methods. A systematic search in four databases (PubMed, Scopus, Web of Science, and ScienceDirect) and Google Scholar was performed via a search strategy using Machine Learning (ML), leukemia, peripheral blood smear (PBS) image, detection, diagnosis, and classification as the keywords. -e average accuracy of the ML methods applied in PBS image analysis to detect leukemia was >97%, indicating that the use of ML could lead to extraordinary outcomes in leukemia detection from PBS images. Among all ML techniques, deep learning (DL) achieved higher precision and sensitivity in detecting different cases of leukemia, compared to its precedents. ML has many applications in analyzing different types of leukemia images, but the use of ML algorithms to detect acute lymphoblastic leukemia (ALL) has attracted the greatest attention in the fields of hematology and artificial intelligence.

[14] Deep learning is a branch of artificial intelligence. In recent years, with the advantages of automatic learning and feature extraction, it has been widely concerned by academic and industrial circles. It has been widely used in image and video processing, voice processing, and natural language processing. At the same time, it has also become a research hotspot in the field of agricultural plant protection, such as plant disease recognition and pest range assessment, etc. The application of deep learning in plant disease recognition can avoid the disadvantages caused by artificial selection of disease spot features, make plant disease feature extraction more objective, and improve the research efficiency and technology transformation speed. This review provides the research progress of deep learning technology in the field of crop leaf disease identification in recent years. In this paper, we present the current trends and challenges for the detection of plant leaf disease using deep learning and advanced imaging techniques. We hope that this work will be a valuable resource for

researchers who study the detection of plant diseases and insect pests. At the same time, we also discussed some of the current challenges and problems that need to be resolved. [15] Identification of the plant diseases is the key to preventing the losses in the yield and quantity of the agricultural product. The studies of the plant diseases mean the studies of visually observable patterns seen on the plant. Health monitoring and disease detection on plant is very critical for sustainable agriculture. It is very difficult to monitor the plant diseases manually. It requires tremendous amount of work, expertise in the plant diseases, and also require the excessive processing time. Hence, image processing is used for the detection of plant diseases. Disease detection involves the steps like image acquisition, image pre-processing, image segmentation, feature extraction and classification. This paper discussed the methods used for the detection of plant diseases using their leaves images.

III. PROPOSED SYSTEM

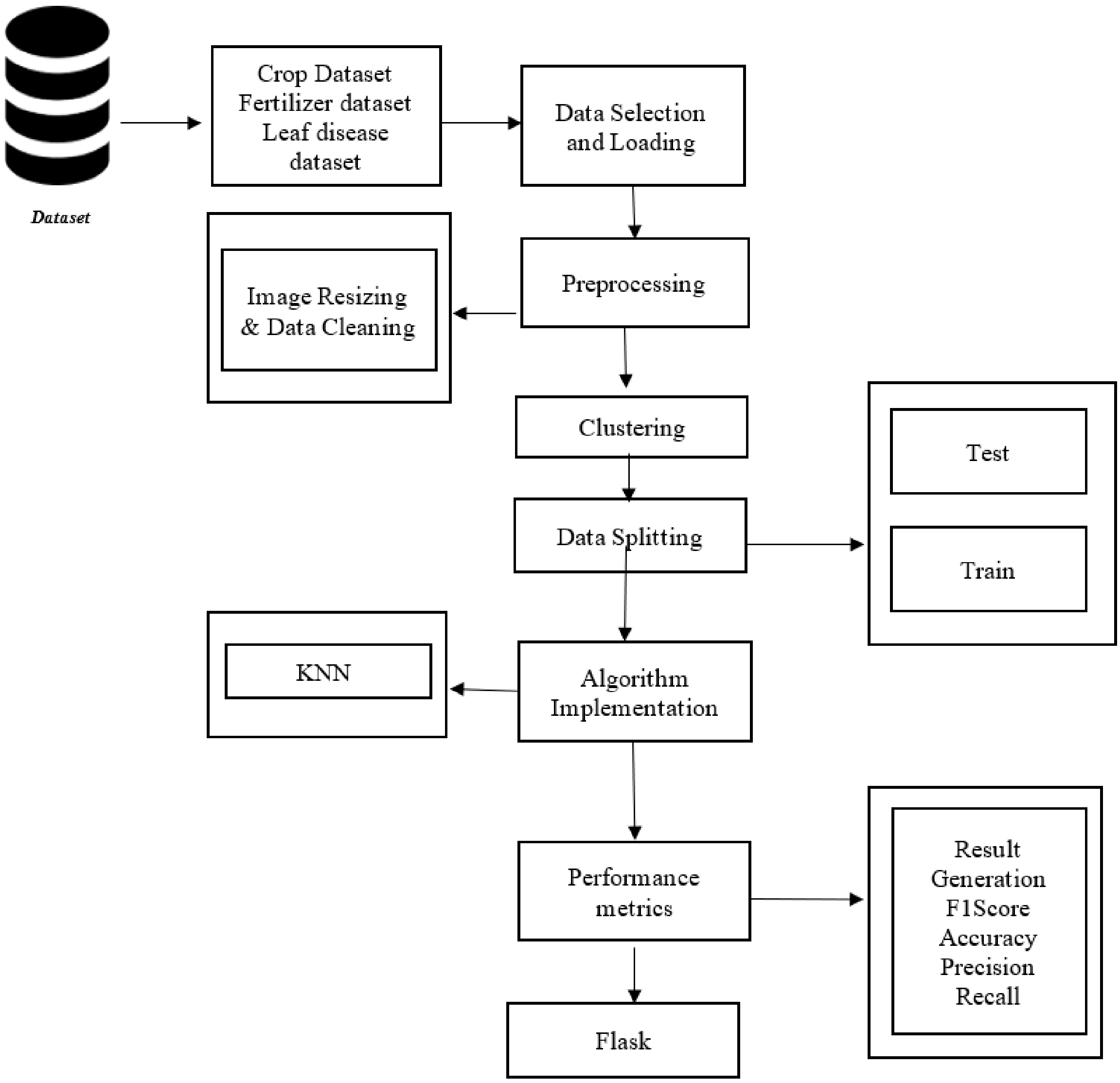

Fig 2: Architectural Diagram of crop yield prediction

Fig 2 shows that, the crop weed dataset was taken as input. The input data was taken from the dataset repository. Then, we have to implement the data pre-processing step. In this step, we have to use image resizing to avoid wrong prediction, and to encode the label for input data. Clustering is used for grouping our similar images. Then, we have to split the dataset into test and train. The data is splitting is based on ratio. In train, most of the data's will be there. In test, smaller portion of the data's will be there. Training portion is used to evaluate the model and testing portion is used to predicting the model. Machine learning such as K- Nearest Neighbor. the experimental results shows that some performance metrics such as accuracy and prediction status. Performance of classifiers is evaluated on the accuracy, precision, recall and F1-score. Result is generated in flask (web application).

The above figure concludes following observations,

1. To create a model that finds to type of crop, fertilizer and leaf disease identification
2. To implement the machine learning algorithm.
3. To implement machine learning like random forest K-nearest neighbor.
4. To enhance the overall performance.

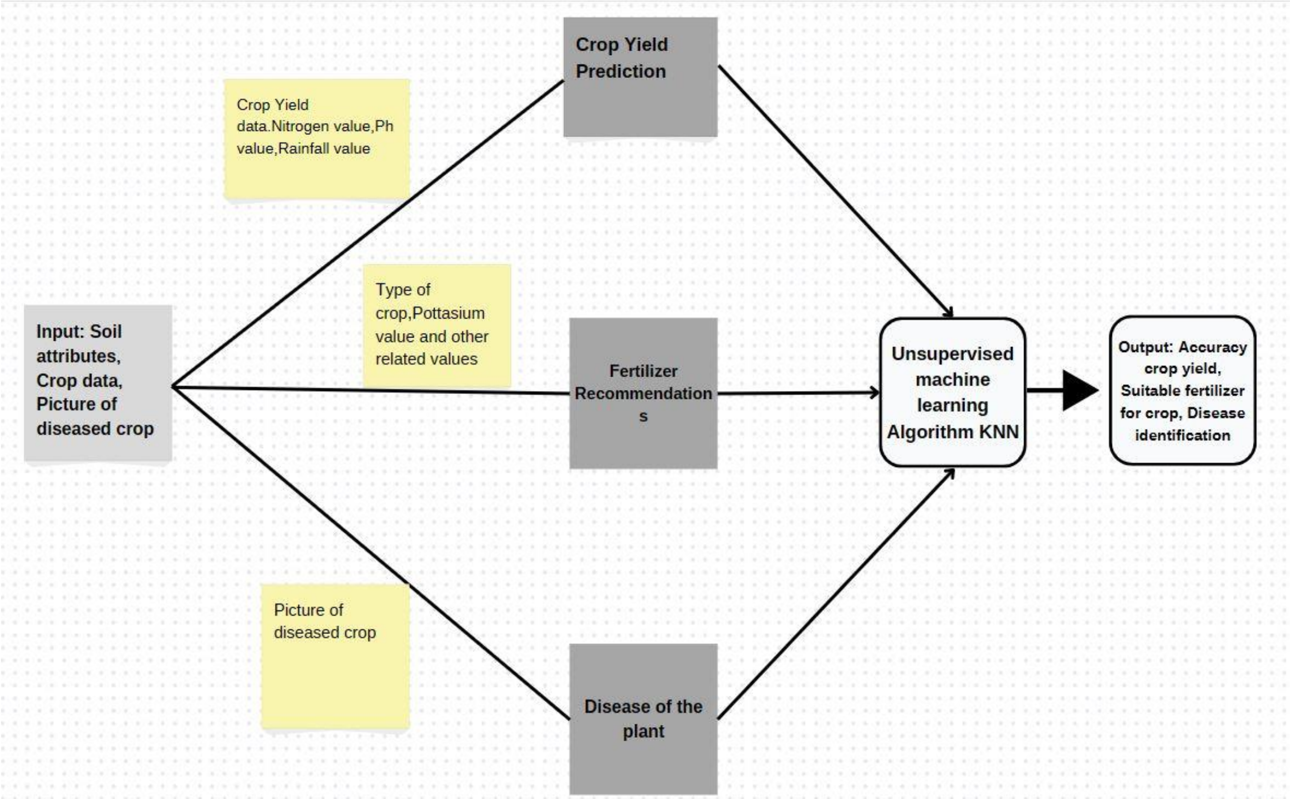

Fig 3: Block Diagram

MODULES IN PROPOSED FRAMEWORK:

In module description we have three modules they are

1. Crop Module
2. Fertilizer Module
3. Disease Module

CROP MODULE:

In this module farmer will give the attributes like pH value, Nitrogen value, Rainfall value and other soil related values. This module will be responsible to give the best suitable crop on the basis of the given soil attribute values. Based on predicted rainfall, soil contents and weather parameters the system will recommend the most suitable crop for cultivation. This system also provides details about required fertilizers like Nitrogen(N), Phosphorus (P) and Potassium (K) in Kg per hectare and display the required seed for a cultivation in Kg per acre for recommended crop. This system as contain some other feature such as display the current market price and approximated yield in quintal per acre for recommended crop. Those all details will helps to farmers for choosing the most profitable crop.

FERTILIZER MODULE:

In this module farmer will give the attributes like type of crop, potassium value and other related values. This module will be responsible to give the best fertilizer suggestions for the given attributes.

DISEASE MODULE:

In this module farmer will upload the image of the diseased plant and this module is responsible to find the disease of the plant and the name of the plant. Farmers will get know the type of disease and name of the plant.

Algorithm:

1. User login to that page and select which module is needed.
2. For crop module.
   a) User give soil attributes
   b) Data is preprocessed to remove unwanted data from the input.
   c) The resulted data is converted into labelled data
   d) By using KNN, performs classification and displayed predicted crop.
3. For fertilizer module.
   a) Fertilizer dataset is collected from dataset repository.
   b) By user inputs it predicts the fertilizer
4. For disease module.
   a) Plant disease dataset is collected from dataset repository.
   b) By user input it classifies the plant disease
   c) Displays the output with other precautionary measures.

IV. CONCLUSION

This paper focuses on the prediction of crop and calculation of its yield with the help of machine learning techniques. Clustering for similar data grouping Several machine learning methodologies used for the calculation of accuracy. KNN classifier was used for the crop prediction for chosen district. Implemented. a system to crop prediction from the collection of past data. The proposed technique helps farmers in decision making of which crop to cultivate in the field. This work is employed to search out the gain knowledge about the crop that can be deployed to make an efficient and useful harvesting. The accurate prediction of different specified crops across different districts will help farmers of India. This improves our Indian economy by maximizing the yield rate of crop.

V. REFERENCE

[1] T. W. Berge, A. H. Aastveit, and H. Fykse, ‘‘Evaluation of an algorithm for automatic detection of broad-leaved weeds in spring cereals,’’ Precis. Agricult., vol. 9, no. 6, pp. 391–405, Dec. 2008.
[2] E. Hamuda, M. Glavin, and E. Jones, ‘‘A survey of image processing techniques for plant extraction and segmentation in the field,’’ Comput. Electron. Agricult., vol. 125, pp. 184–199, Jul. 2016.
[3] H. Mennan, K. Jabran, B. H. Zandstra, and F. Pala, ‘‘Non-chemical weed management in vegetables by using cover crops: A review,’’ Agronomy, vol. 10, no. 2, p. 257, Feb. 2020.
[4] X. Dai, Y. Xu, J. Zheng, and H. Song, ‘‘Analysis of the variability of pesticide concentration downstream of inline mixers for direct nozzle injection systems,’’ Biosyst. Eng., vol. 180, pp. 59–69, Apr. 2019.
[5] D. C. Slaughter, D. K. Giles, and D. Downey, ‘‘Autonomous robotic weed control systems: A review,’’ Comput. Electron. Agricult., vol. 61, no. 1, pp. 63–78, Apr. 2008. [6] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, ‘‘Machine learning in agriculture: A review,’’ Sensors, vol. 18, no. 8, p. 2674, Aug. 2018.
[6] K. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, ‘‘Machine learning in agriculture: A review,’’ Sensors, vol. 18, no. 8, p. 2674, Aug. 2018.
[7]A. Wang, W. Zhang, and X. Wei, ‘‘A review on weed detection using ground-based machine vision and image processing techniques,’’Comput. Electron. Agricult., vol. 158, pp. 226–240, Mar. 2019.

[8] F. Ahmed, H. A. Al-Mamun, A. S. M. H. Bari, E. Hossain, and P. Kwan, ‘‘Classification of crops and weeds from digital images: A support vector machine approach,’’ Crop Protection, vol. 40, pp. 98–104, Oct. 2012.
[9] P. Herrera, J. Dorado, and Á. Ribeiro, ‘‘A novel approach for weed type classification based on shape descriptors and a fuzzy decision-making method,’’ Sensors, vol. 14, no. 8, pp. 15304–15324, Aug. 2014.
[10] Y. Chen, X. Jin, L. Tang, J. Che, Y. Sun, and J. Chen, ‘‘Intra-row weed recognitionusingplantspacinginformationinstereoimages,’’presentedat the Kansas City, Missouri, St.Joseph,MI,USA,2013.[Online].Available: http://elibrary.asabe.org/abstract.asp?aid=43357&t=5
[11] S. Shafique and S. Tehsin, ‘‘Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks,’’ Technol. Cancer Res. Treatment, vol. 17, Sep. 2018, Art. no. 1533033818802789.

[12] B. K. Das and H. S. Dutta, ‘‘Infection level identification for leukemia detection using optimized support vector neural network,’’ Imag. Sci. J., vol. 67, no. 8, pp. 417–433, Nov. 2019.

[13] K. K. Jha, P. Das, and H. S. Dutta, ‘‘FAB classification based leukemia identification and prediction using machine learning,’’ in Proc. Int. Conf. Syst., Comput., Autom. Netw. (ICSCAN), Jul. 2020, pp. 1–6.

[14] Li, L., Zhang, S., & Wang, B. (2021). Plant Disease Detection and Classification by Deep Learning—A Review. IEEE Access, 9, 56683–56698. doi:10.1109/access.2021.3069646
10.1109/access.2021.3069646

[15] Sachin D. Khirade, A. B. Patil, 2015 International Conference on Computing Communication Control and Automation.