International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 04 Issue: 05 | May -2020 ISSN: 2582-3930

CROP YIELD PREDICTION USING MACHINE LEARNING

P.K.Akulwar[1],Pooja Sachin Jirage[2], Pratiksha Rajendra Patil[3] , Sarika Surendra Mali[4]

Snehal Sunil Kandekari[5], Mayuri Prakah Koshti[6]

[1]*Professor, Department of Computer Science and Engineering, Sanjay Ghodawat Institute, Atigre*
[2]*Student, Department of Computer Science and Engineering, Sanjay Ghodawat Institute, Atigre*
[3]*Student, Department of Computer Science and Engineering, Sanjay Ghodawat Institute, Atigre*
[4]*Student, Department of Computer Science and Engineering, Sanjay Ghodawat Institute, Atigre*
[5]*Student, Department of Computer Science and Engineering, Sanjay Ghodawat Institute, Atigre*
[6]*Student, Department of Computer Science and Engineering, Sanjay Ghodawat Institute, Atigre*

Abstract - Agricultural planning is most important part to estimate a future agricultural production. Achieving maximum rate of crop yield using limited land resource is an important challenge for farmers. Agriculture has the largest contribution in the GDP(Gross Domestic Product) of our country. But still the farmer's don't get worth price of the crops. It mostly happens due to improper irrigation or inappropriate crops selection or also sometimes the crop yield is less than that of expected. By analyzing the soil and atmosphere at particular region in order to achieve more crop yield and the net crop yield can be predict. This prediction will help the farmers to choose appropriate crops for their farm according to the soil type, temperature, humidity, water level, spacing depth, soil PH, season, fertilizer and months. This prediction can be carried out using Random Forest and other machine learning classification algorithms.

Key Words: Crop yield, Machine Learning, Prediction, Convolutional Neural Network, Linear Regression.

1. INTRODUCTION

Agriculture plays important part of Indian economy. Achieving maximum rate of crop yield using limited land resource is a goal of agricultural planning in an agro-based Country. Selection of crop is important issue in agricultural planning.

There is always a significant risk factor for the farmers to decide which particular crop should grow during particular season, on Particular land resource. It depends on different parameters such as production rate, cost and different government policies. Crop selector method could be applicable for minimize losses of crop when unfavorable conditions occurs and this selector could be used to maximize the crop yield rate when potential exists for favorable growing conditions. Irrespective of the capital put in terms of soil, water and quality of seeds of the crop rate production the crop may fail bringing disastrous losses to the farmer and his family.

There is significant research in agriculture, by which it helps to attempt the improvement of crop production to help the farmers with the help of machine learning techniques. It requires dataset available from past experiences of farming a particular crop in particular season. A statistical and machine learning both techniques were modeled. Crop production rate depends on geography of a region, weather condition, soil type and harvesting methods. Different subsets of these influencing parameters are used for different crops by different prediction models. Machine learning technique is used for prediction of crop yield using different algorithms. Machine learning methods which are widely used in prediction technique are regression tree, random forest, convolution neural network and K-nearest algorithm.

2. PROBLEM STATEMENT

Crop production rate depends on geographical region, weather condition soil type, soil composition potassium, organic carbon, calcium, and harvesting methods. Traditionally, monitoring techniques does not gather the crop conditions properly and prediction results were not yet optimized. Therefore to overcome this problem alternate way is to design a system that will identify the crop type and display the disease of that crop (i.e. Gray Mold or Rhizopus), which will help to predict crop yield.

3. LITERATURE SURVEY

The author [1]Mahabadi et alhas described, Artificial Neural Network models that are ready with varied variety of neurons in hidden layer, back dissemination learning calculations. Modifying these parameters inspired the capability of the system to create up a perfect capability to foresee crop yield. The educational rate and diversity of hidden nodes for the most part have an effect on show conduct. For the foremost half, less hidden nodes were needed because the quantity of knowledge diminished. The most effective models have less hidden nodes than the start variety of nodes. ANN models with a lot of nodes could have resulted in over fitting as hostile learning their connections. RMSE (Root mean sq. error), was used to assess the execution of created model.

The author [2] Hemageethaa has described, Naïve Bayes algorithm which is used for yield prediction. Focuses mainly on different soil parameters like pH, Nitrogen, moisture etc and comparison accuracy is also presented. Only 77% of accuracy is achieved. Hemageethaa mainly focuses on using the soil parameters like pH, Nitrogen, moisture etc for crop yield prediction. Naive Bayes algorithm is used to classify the soil and 77% accuracy is achieved. Appropriate algorithm is used to associate the soil with the crops that could provide maximum yield in them. A comparison of exactly achieved during classification using Naïve Bayes, J48 and JRIP is also presented. Sometimes farmers are not aware about the crop which suits their soil quality, soil nutrients and soil composition. The work proposes to help farmers check the soil quality depending on the analysis done based on data mining approach. Thus the system focuses on checking the soil quality to predict the crop suitable for cultivation according to their soil type and maximize the crop yield with recommending appropriate fertilizer.

The author [3] Grajales has developed web application using Open source tools. The details of the selected location from map are available at one look to the user. Grajales et al have proposed a web application that utilizes open dataset like historical production, land cover, local climate conditions and integrates them to provide easy access to the farmers. The proposed architecture mainly focuses on open source tools for the development of the application. The user can select location from map for which the details are available at one click.

The author [4] Sujatha has described, Naïve Bayes, J48, random forests, support vector machines, and artificial neural networks are implemented. Climate data and Crop parameters are mainly used for crop yield is predicted. Other parameters like soil are not considered. Sujatha describes the purpose of various classification techniques that could be used for crop yield prediction. A few of the data mining methods, such as the Naïve Bayes, J48, random forests, support vector machines, artificial neural networks were presented. A system using climate data and crop parameters are used to predict crop growth has been proposed.

The author [5] Sellam et al has described, Regression Analysis (RA), Linear Regression (LR) describes the environmental factors that affect the crop yield and the relationship among these parameters is also established. The different environmental parameters like the Area under Cultivation, Annual Rainfall and Food Price Index that influences the yield of crop and the relationship among these parameters is established. Using Regression Analysis (RA), Linear Regression (LR) the different environmental factors and their infliction on crop yield is analyzed.

4. PROPOSED SYSTEM

The Raspberry Pi's camera will capture the image. Then this image is processed and features of the image. (like Shape, Color, position etc) are extracted and the crop type is identified. Using Convolution Neural Network Algorithm, predict the crop disease. Using Linear Regression algorithm, Crop Selection Method is performed. By using Convolution Neural Network (CNN) algorithm system will identify the crop type. CNN algorithm is used to classify the image by its features.

Diseases were identified manually and based on experience. Plants were protected by using chemicals and fertilizers without analyzing quantity of chemicals, fertilizers or pesticides needed for particular crop. After identifying the crop type the System will display the disease of that crop in percentage. Crop selection method refers to a method of selecting crop over a particular season depending upon various environmental as well as economic factors for the maximum advantages. These factors are rainfall levels, average temperature, soil type, market prices and demand etc. Using Linear Regression Algorithm the infected crop is removed and user will get the particular area of land. Based on the result the largest probability produced is taken for projection. The correctness can be predicted by comparing the resultant class value with the test data set. The accuracy can range from 0% to 100%. After performing above modules user will get the exact yield of crop.

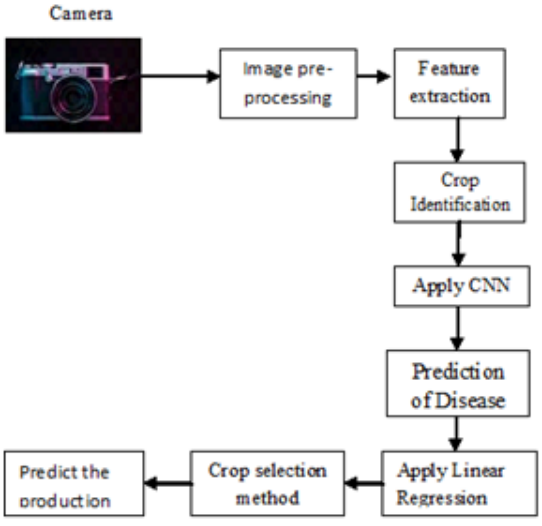

Fig -1: System Architecture

Raspberry Pi4

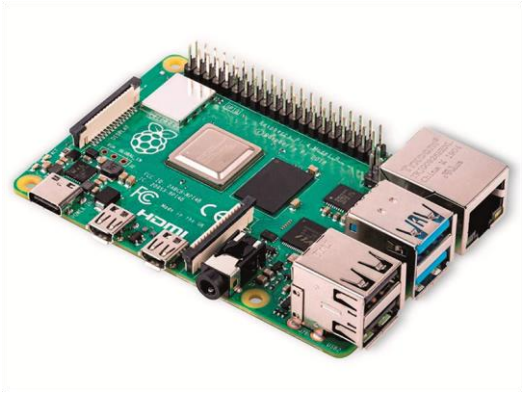

Fig -2: Raspberry Pi4

The Camera Module can be used to take high-definition video, as well as stills pictures. It's easy to use for starters, but has plenty to offer advanced users if you're looking to expand your knowledge. There are lots of examples online of people using it for period-lapse, slow-motion, and other video cleverness. You can also use the libraries we bunch with the camera to create effects. You can read all the gory details about IMX219 as well as the Exmor R back- illuminated sensor architecture on Sony's website, but sufficient to say this is more than just a resolution upgrade: it's a jump forward in image quality, colour fidelity, and low-light performance. It supports 1080p30, 720p60 as well as VGA90 video modes, and still captures. It is attached to the CSI (Camera Serial Interface) port via a 15 cm ribbon cable on the Raspberry Pi. The camera working with all models of Raspberry Pi 1, 2, 3 and 4. It can be gain access via the MMAL and V4L APIs, and there are numerous third-party libraries built for it, including the Picamera Python library. See the Getting Started with Picamera resource to study how to use it. The camera module is very famous in home security applications, and in wildlife camera traps.

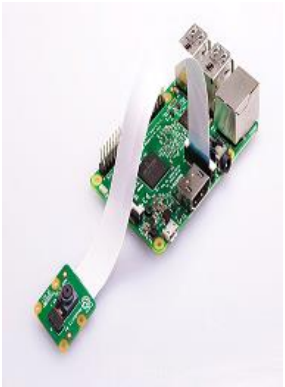

Fig -3: Camera Module

5. ALGORITHM

1. Convolution Neural Network A Convolution Neural Network (CNN) is deep learning algorithm which takes input images to various aspects/objects in the image and differentiate them one from another. CNN is machine learning algorithm for machines to understand the features of the image with foresight and remember the features to guess whether the name of the new image fed to the machine. The algorithm CNN takes image as an input and identify the image i.e. whether the image is fruit, flower, animal etc then classify that image under certain categories like if CNN algorithm detects the object id Animal then it will classify it whether it is dog, cat, lion etc. the input images are array of pixels and it depends on image resolution.

Training neural network: Data-set will be generated and stored accordingly in the folders. Images are collected with readily available browser extensions for bulk downloads. These images are then preprocessed the same manner on the computer as to how they are going to be processed in the hand held device. The images are then feed to the neural network using python- tensorflow or python-scikitlearn implementation. Readily available and proven effective ANN model provided by Google's open source project is selected for the neural network. The last layer of the neural net- work is then trained to fit our specific need of detection diseases in selected crops. The training is done for 30 epochs and then the accuracy is checked using tensor- Flow library functions.

2. Crop Selection Method

Step 1: place the camera on raspberry pi's port

Step 2: the input for the camera is crop so put the crop for e.g. Strawberry in front of camera. Camera will capture the image.

Step 3: classify the image and identify the disease.

```
# classify the input image

print("[INFO] classifying image...")

proba = model.predict(image)[0]

idxs = np.argsort(proba)[::-1][:2]
```

Step 4: Display the disease type in percentage wise from 0-100%.

Step 5: Calculate the crop count and display the final crop yield prediction.

6. RESULT AND DISCUSSION

To identify crop yield prediction the system will use the machine learning algorithms. First it will identify the crop (i.e. if we take an example strawberry fruit) the algorithm recognize the fruit then it will identify disease on the crop if the crop is infected it will display the disease in percentage manner from 0-100%. Then it will calculate the production of crop by removing the infected crop. At the end user will get the exact crop yield production.

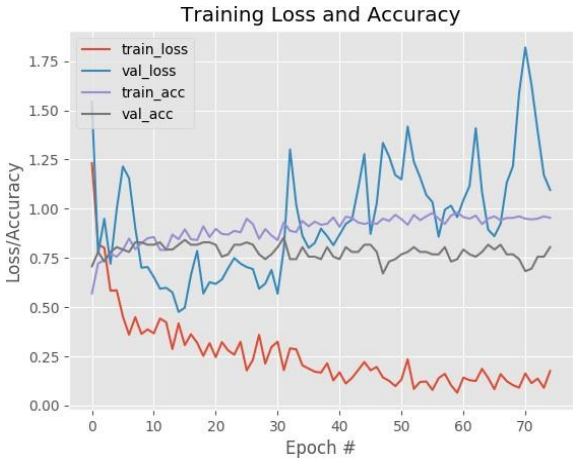

Chart 1: Shows the training loss and accuracy of model.

7. CONCLUSIONS

Agriculture is the main resource of economy Now a day's disease occurring on crops, has increased which has reduce the productivity of the crop from 0-100% percent. Manually identifying the disease is the difficult and time consuming task and it will not get worth price. Instead of manually identifying disease image processing and machine learning technique can be used to identify the disease which is less time consuming, cost effective and more accurate the manual identification technique. Machine learning technique can be implemented using different programming such as R programming, Python, Java, etc. the system uses python programming language to implement machine learning approaches. This project helps in understanding the creation of an System that will do image processing and identify the crop diseases and predict the crop yield production using machine learning approach.

REFERENCES

[1] Rakesh Kumar, M.P. Singh, Prabhat Kumar and J.P. Singh."Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique",International Journal of Innovative Research in Computer and Communication Engineering, 2010.

[2] J.RamirezVillegas and A.Challinor, "Assessing relevant climate data for agricultural applications",Agricultural Forest Meteorology, 2012.

[3] Pushpa Mohan1, Dr. Kiran Kumari Patil," Survey on Crop and Weather Forecasting based on Agriculture related Statistical Data ",International Journal of Innovative Research in Computer and Communication Engineering, 2007.

[4] Kodimalar Palanivel, Chellammal Surianarayanan, "An Approach for Prediction of Crop Yield Using Machine Learning", International Journal of Computer Engineering and Technology, 2019.

[5] Hemantkumar Wani,Nilima Ashtankar, " An appropriate model predicting diseases of crops using machine learning algorithms", International Conference on Advanced Computing and Communication Systems (ICACCS), 2017.

[6] Prof. D.S. Zingade1,Omkar Buchade2,Nilesh Mehta3,Shubham Ghodekar4,Chandan Mehta," Crop Prediction System using Machine Learning", International Journal of Advance Engineering and Research Development,Dec.-2017

[7] R.Nagini, Dr. T.V. Rajnikanth, B.V. Kiranmayee, "Agriculture Yield Prediction Using Predictive Analytic Techniques", 2nd International Conference on Contemporary Computing and Informatics, 2016.

[8] Awanit Kumar, Shiv Kumar, "Prediction of production of crops using K-Means and Fuzzy Logic",IJCSMC, 2015.

[9] S. Veenadhari , Bharat Misra , CD Singh," Machine learning approach for forecasting cropyield based on climatic parameters ",International Conference on Computer Communication and Informatics, 2014.

[10] Sujatha, R., Isakki, P., "A study on crop yield forecasting using classification techniques", International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE), 2016.