JuniKhyat (UGC Care Group I Listed Journal)
ISSN: 2278-463 Vol-13 Issue-01 May 2023

CROP HARVESTING FORECAST AND DISEASE SENSING ALGORITHM BY EMPLOYING MACHINE LEARNING TECHNIQUES

Dr.M.Ajay Kumar[1]., S.Pragathi[2].,L.Anusha[3].,A.Bhoomika[4].,B.Yamuna[5]

[1]Associate Professor, Department Of ECE., Malla Reddy College Of Engineering For Women., Maisammaguda.,Medchal., Ts, India (ajayajju126@gmail.com)

[2, 3, 4,5]B.TechECE, (19RG1A0451, 19RG1A0429, 20RG1A0401,20RG1A0403),

Malla Reddy College Of Engineering For Women., Maisammaguda., Medchal., Ts, India

Abstract:

India's agricultural sector is vital to the country's overall economy. About 18% of India's GDP comes from the agricultural sector. There has been a significant decline in agricultural activity in India during the last several years. Climate change and the use of crude technologies are the primary drivers of these dramatic shifts. Farmers might benefit from monitoring their crops and making predictions about their yield far in advance of harvest. For many issues in agriculture, machine learning is crucial to finding workable and successful solutions. In this study, we explore the use of machine learning algorithms for the diagnosis of crop diseases and the prediction of their effects on production. The model is constructed using the convolutional neural network (CNN) technique and the built-in image processing capability. TensorFlow illness detection makes use of this model. The yield prediction model is updated by constructing a second model using the Random Forest Algorithm.

Keywords:

Disease Detection, Yield Prediction, Machine Learning, Convolution neural network (CNN), Random Forest Regression, Image processing, User Interface Web Application.

Introduction

India's agricultural sector is crucial to the country's economy, supporting between 50 and 60 percent of the country's workers. More than 60 percent of domestic consumption is met by agricultural output alone, making agriculture the primary source of food supply for a sizable portion of our country's population. There is a vast variety of vegetable and fruit crops from which farmers may choose. However, farmers take a big chance every time they choose to plant a specific crop at a specific time on specific land. Growing these plants for maximum output and quality is, however, a complex technological process. In spite of the farmer's best efforts, his harvest may fail, resulting in catastrophic losses for him and his family and maybe even leading to bankruptcy and suicide. Since they still practice traditional farming methods, they have little idea of how to solve this issue. If the farmer could predict the crop's production, he could make an informed decision about what to grow. After deciding what to grow, it is the farmer's obligation to keep an eye on his fields for signs of plant disease, which may have a major impact on the harvest. The research project's software will have a straightforward interface that will help farmers make informed decisions about crop output and disease prevention and treatment. Machine learning is the most effective method for solving such agricultural issues. Algorithms are used in Machine Learning to examine data and draw conclusions. The information is then applied to a global problem in need of a solution. Create our app, we used supervised machine learning to construct a model that is trained using historical crop yield data and several picture datasets of diseases affecting the target crop. This learned model is then used to the input data to make predictions about yield and illness.

Methodology

A. Random Forest Regression

For both regression and classification purposes, the random forest technique employs an ensemble learning approach. Supervised Learning describes this method of data analysis. The Ensemble method is a strategy that uses the combined forecasts of many different machine learning algorithms. When compared to a single model, ensemble models produce more reliable forecasts. Bootstrap and Boosting Aggregation are two of the several forms of learning. Instead of being a boosting method, random forest is a bagging method. In bagging, all models are executed separately before their results are combined without favouring any one model over the others. During training, the bagging model creates a forest of decision trees from which to make predictions.

Input from many decision tree classifiers may be pooled to form a random forest ensemble. This information is combined into an ensemble model by model voting or averaging. Some tweaks that aid in this procedure are: First, there is a cap on the proportion of features that can be used to partition each node. This guarantees that the ensemble model does not over-rely on any one feature while yet making full use of all available predictors. Overfitting may be avoided since each tree uses a random sample of the original data to generate its splits. Random Forest's primary benefits and characteristics are:

1. It is the most accurate learning algorithm available.

2. It produces a highly accurate classifier.

3. It runs efficiently on large databases.

4. Without variable deletion it can handle thousands of input variables. 5. Highly effective method for estimating missing data.

6. When a sizeable proportion of the data is missing it maintains accuracy.

Neural Network with Convolutions (B) Convolutional neural networks (Convent’s) are a kind of Deep Learning algorithm that can analyse an input picture and classify its contents based on the weights and biases it is taught. The design of the Convent is similar to the way in which neurons in the human brain communicate with one another. The structure of the visual cortex served as inspiration. CNN typically operates by transmitting an image to a network that is reading in the picture. The input picture then through a convolutional stage of the network, which consists of an unlimited number of steps. At last, a neural network allows us to foretell the image's result.

Four main components of a CNN are:

1. Convolution 2. Non-Linearity (REL)

3. Pooling or Sub Sampling

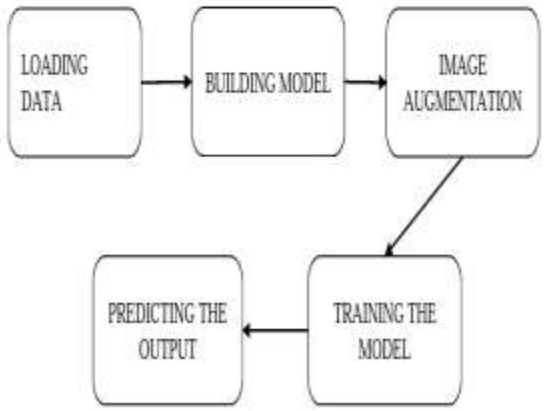
4. Classification (Fully Connected Layer)

After the input image has been processed by each of the aforementioned layers, it is time to generate the final output. The first layer, called the Convolution layer, is responsible for extracting characteristics from the picture data. The filter squares are used in convolution to learn visual attributes and preserve the pixel-to-pixel connection. Its values are multiplied by the pixels' original values when the filter is applied. All of these calculations are multiplied together afterwards. One last digit is received. A matrix is produced;however, it is less than the input matrix after performing the filter through all locations. By convolving a picture with several filters, it is possible to perform operations like detecting edges, blurring, and sharpening. The nonlinear layer is accompanied by the pooling layer. The image's width and height are used in a down sampling process. This results in a smaller overall picture volume. Data from convolutional layers is sent into the fully connected layer. An N-dimensional vector, where N is the number of classes from which the model picks the required class, is the outcome of adding a completely linked layer to the end of the network.

Implementation

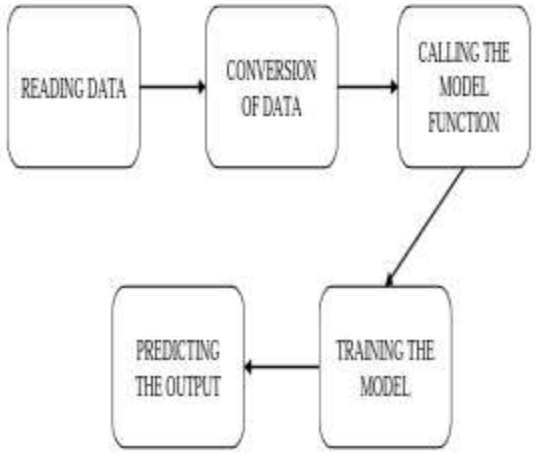
To facilitate communication between the system and the farmer, we have designed a simple web application. The online app is broken down even further into its two primary sections, which are the identification of diseases and the forecasting of agricultural yields. The user is first prompted to choose a crop by the disease detection department before being sent to a website where a picture of the affected area of that crop may be uploaded. Once the image has been uploaded, it is sent to the model, where the disease is identified, and its treatment options are displayed. The user is first prompted to input the region, growing season, and crop type before proceeding to the yield forecast section. Once the user has provided all of this data, it is sent to the model, which then makes a yield prediction based on the inputs it has been given. The website then shows the expected value.

Disease Detection:



The accompanying graphic depicts the process flow for the illness detection section. An '. nay' file is created by loading a list containing picture datasets of certain plants; this list is then turned into a NumPy array. The CNN model is developed utilizing the native capabilities of TensorFlow and Kera's. Convolutional, activation, pooling, and fully linked layers make up this model. After some arithmetic processing, thegnly files are saved and given along to the ImageDataGenerator function, which is responsible for the actual picture augmentation. The result of this function is fed into a previously constructed CNN model, which is then stored. The illness prediction for the input is made by loading the previously stored model.

Yield Prediction:

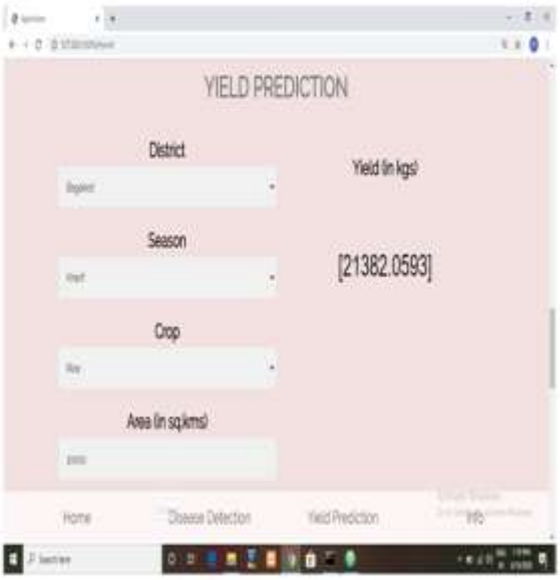


The following illustration depicts a method for forecasting crops. Datasets stored in comma-separated values (CSV) files are read and linked to a target variable. After that, a NumPy array is created to store the data. The Random Forest Regressor model function is available after the skarn package has been loaded. NumPy array values are stored as variables, and the model is trained by using the fit function to those variables. After the training is complete, the model is written to disk in the form of a pickle file. The outcome of a given input is predicted using this model.

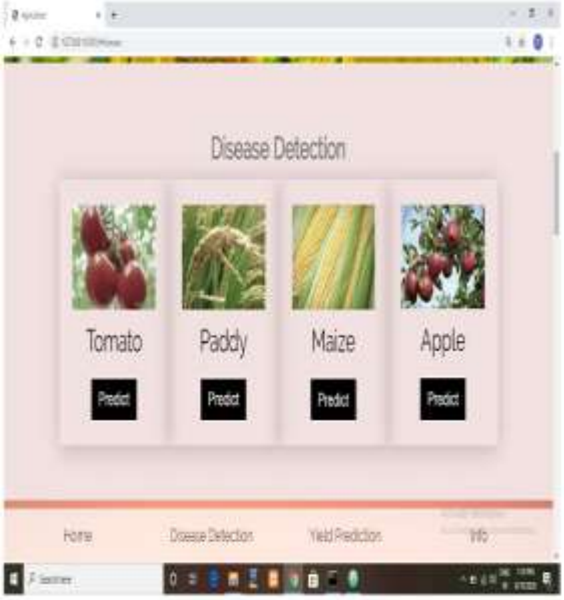
Result

The study's overarching objective was to develop a simple online tool for agriculturists to utilize in order to forecast crop yields and identify afflictions in advance of harvest. Django was used to turn the system's backend python code into a web page displaying the system's output.

The screenshot above depicts the web app's landing page, which invites users to anticipate yield, identify sickness, and provide comments.



The yield forecasting mini-web app is shown below. The user is prompted to choose the crop, the district, the season, and the location. The web app then shows the expected return on investment (ROI) based on the user's input.



The following diagram depicts the sub-web app that allows the user to choose a specific crop for disease diagnosis. When the crop is chosen, the user is taken to the subsequent page.

The user is prompted to provide a photograph of the affected area of the leaf; the system then analyses the input to identify the illness and provide treatment options.

Conclusion

The model used to create this web app has a 90% confidence interval. In the yield prediction section, we take into account datasets of previous year's crop production, which are generated by thinking about things like the season, the crop, and the rainfall at a given region. However, this does not take into account all of the variables that could affect the outcome of the experiment, for example, soil fertility and pests. Only a subset of the crops grown in Karnataka are included in our disease prediction model. These include tomato, sugarcane, rice, and others. Additional crop diseases can be predicted with this method.

References

1. ***Mrs. Shruthi U, Craigavon V, Raghavendra B K, "A Review on Machine Learning Classification Techniques for Plant Disease Detection," ICACCS, IEEE, 2019.***
2. ***M. P. Vaishnavi, K. Suganya Devi, P. Srinivasan, G. ArutPerumJothi, "Detection and Classification of Groundnut Leaf Diseases using KNN classifier," IEEE, 2019.***
3. ***L. Sherly Puspha Annabel, "Machine Learning for Plant Leaf Disease Detection and Classification – A Review," International Conference on Communication and Signal Processing, IEEE, 2019.***
4. ***Gurleen Kaur Sandhu, Rajbir Kaur, "Plant Disease Detection Techniques: A Review," ICACTM, IEEE, 2019.***
5. ***Md. Selim Hossain, Rokeya MuthannaMoue, "Recognition and Detection of Tea Leaf's Diseases Using Support Vector Machine," International Colloquium on Signal Processing & its Applications, IEEE, 2018.***
6. ***Monaural Islam, Anh Dinh, Khan Wahid, "Detection of Potato Diseases Using Image Segmentation and Multiclass Support Vector Machine," CCECE, IEEE, 2017.***
7. ***Jobin Francis, Anto Sahaya Dahs D, Anoop B K, "Identification of Leaf Diseases in Pepper Plants using Soft Computing Techniques," IEEE, 2016.***
8. ***Vijai Singh, Varsha, "Detection of the unhealthy region of plant leaves using Image Processing and Genetic Algorithm," ICACEA, IEEE, 2015.***
9. ***Mrunmayee Dhak ate, Ingole A. B., "Diagnosis of Pomegranate Plant Diseases using Neural Network," IEEE, 2015.***
10. ***Sachin D. Charade, A. B. Patil, "Plant Disease Detection Using Image Processing," IEEE, 2015.***