International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8354
***Article Received:** 15 July 2023 **Revised:** 06 September 2023 **Accepted:** 19 September 2023*

Sustainable Agriculture Practice using Machine Learning

P. Ashwini[1], K. Srividya[2], Dr. N. Vadivelan[3]
[1]Department of CSE, VASAVI College of Engineering
Hyderbad, India
ashwinireddy90@gmail.com
[2]Department of CSE, VASAVI college of Engineering
Hyderbad, India
ksrividya0508@gmail.com
[3]Professor
[3]Department of CSE, Teegala Krishna Reddy Engineering College
Hyderbad, India

Abstract- The changing climate has caused unpredictable rainfall, unusual temperature drops, and heat waves, leading to considerable damage to the environment. Fortunately Machine Learning has provided effective tools to address global issues, including agriculture. By employing different ML algorithms, it is possible to solve the agricultural problems caused by these climate changes. The objective of this article is to develop a system for crop recommendation and disease detection in a plant. Publicly available datasets were used for both tasks. For the crop recommendation system, feature extraction was performed, and the dataset was trained using various Machine Learning algorithms, namely Decision Tree, Logistic Regression, Random Forest, Support Vector Machine (SVM) and Multilayer Perceptron. The random forest algorithm achieved an excellent accuracy of 99.31%.For the plant disease identification system, CNN architectures like - VGG16, ResNet50, and EfficientNetV2 - were trained and compared. Among these, EfficientNetV2 achieved high accuracy of 96.07%.

Keywords- SVM, Multilayer Perceptron(MLP), Random Forest(RF), CNN, VGG16, ResNet50, EfficientNetV2.

I. INTRODUCTION

Machines with the ability to learn have the potential to tackle difficult problems that are difficult for people to solve. It is applicable in various fields, including agriculture, sports and business. It has the potential to carry out activities including categorization, prediction, and identification. The fundamental purpose of this article is to develop a website that addresses two urgent challenges, namely crop recommendation and crop disease identification. This approach will cater to the needs of the agricultural industry and will cater to the needs of the agricultural industry [1].In order to come up with answers to these problems, models were trained with the help of datasets that were available to the general public, and their findings were compared. The models demonstrated an acceptable level of accuracy was included and it may be utilized in the cloud. This was done so in order to make the models more accessible.

Over the course of the past five years, the climate has undergone substantial shifts, which has had a huge influence on agriculture. Lack of information of scientific agricultural practices frequently leads to the selection of crops that are not suitable for the intended use [2]. When it comes to making decisions, farmers sometimes have to rely on a limited amount of experience, which can make them more prone to making mistakes. As a direct result of this, the agricultural industry suffers enormous losses as a direct result of the inefficient exploitation of essential information, such as the composition of the soil, the pH of the soil, and the prompt detection of plant diseases [3]. As a direct result of this, the agricultural sector suffers enormous losses as a direct result of the inefficient exploitation of vital information. The issue can be fixed by employing modern technology in an effective manner, which will make this possible. Due to the fact that it employs both machine learning and the web, this strategy is able to target individuals who have access to mobile phones that are capable of connecting to the internet.This paper is arranged in 8 sections; Section 2 presents Literature review, Section 3&4 covers Classification algorithms, Section 5&6 describes crop recommendation system and plant disease detection, Section 7 presents Results and section 8 includes conclusion & future scope.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8354
***Article Received:** 15 July 2023 **Revised:** 06 September 2023 **Accepted:** 19 September 2023*

II. LITERATURE SURVEY

Efforts have been made to address the challenges of plant disease detection and recommendation of crop. G Chauhan and A. Chaudhary has suggested a type of crop based on soil. They utilized Decision Trees and Random Forest to make predictions, and the results demonstrated the potential of using a random forest classifier to predict crops based on land patterns [4]. For plant disease identification, S.P. Mohanty used a dataset along with their grayscale and segmented part. The paper published by him talks about the classification task done by using AlexNet. Jose M. Cadenas et.al presented a decision support system based on time series datasets and inference engine to make decisions in agriculture field with intelligence [5].Yukimasa Kaneda et.al implemented a multi model SW-SVR and DNN approach for predicting plant water stress. The proposed method predicts water stress of plants based on environmental and plant image dataset more accurately [6].Sharada P. Mohanty et.al proposed deep learning model for image based plant dieses detection. In this model authors used public data set which contains 54,306 samples of healthy and dieses plant leaves and it achieved an accuracy of 99.35% [7].

III. CLASSIFICATION ALGORITHM

This section gives an overview of the ML classification algorithms used for plant dieses classification.

A. Logistic Regression

Logistic regression is similar to Linear Regression, but it is used to solve classification tasks. The function used in the logistic regression is a sigmoid function which gives a value between 0 and 1.

B. Decision Tree

Decision trees are versatile tools used in classification and regression problems. It is represented like a tree structure where leaves represent final decisions. Decision Tree rely on entropy, which measures the uncertainty or disorder in a dataset. Information gain, on the other hand, measures the reduction of uncertainty provided by a specific feature and is a crucial factor in determining which attribute should be chosen as the decision node or root node [8].

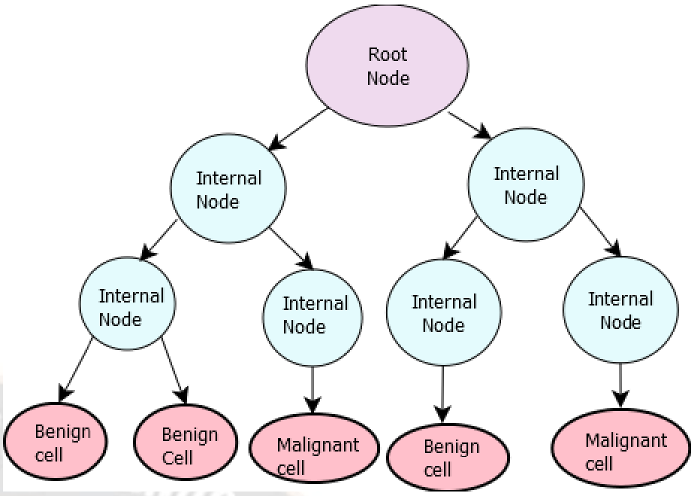

Fig 2: Decision Tree for crop recommendation System

C. Support Vector Machine

SVM is a supervised model of ML and it is most frequently used for the classification of various datasets, despite the fact that it is capable of handling regression issues as well. The SVM algorithm consists of n-dimensional space which includes every data item is a data point. Where n represents a number of features and n-dimensional space corresponds to the number of dimensions. Each feature value indicates a particular coordinate. The next step in the classification process involves locating the hyper-plane that most clearly demarcates the difference between the two classes [9].

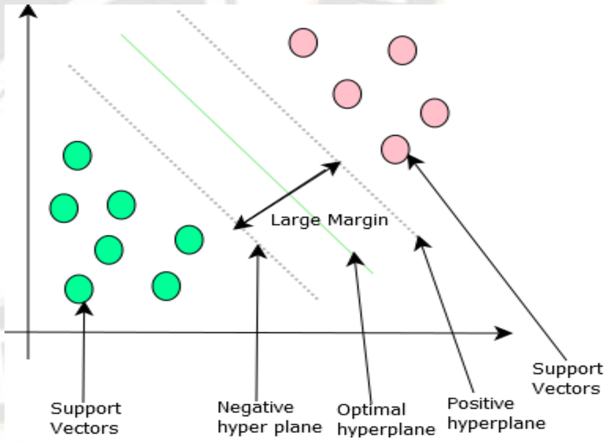

Fig 1: Support Vector Machine for crop recommendation System

D. Multi-Layer Perceptron

The multilayer perceptron (MLP) neural network is a classifier to carry out the tasks connected with classification. The MLP Classifier is a classification algorithm that, unlike other classification algorithms such as Support Vectors and the Naive Bayes Classifier, relies on a neural network to carry out the classification. Other classification techniques include the Naive Bayes Classifier and the Support Vector Classifier [9]. Support vectors and the Naive Bayes classifier are two examples of these

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8354
***Article Received:** 15 July 2023 **Revised:** 06 September 2023 **Accepted:** 19 September 2023*

additional classification techniques. An artificial neural network that makes use of feed forward in order to turn input data sets into output and is referred to as a Multilayer Perceptron, which is also sometimes abbreviated to MLP for short.

E. Random Forest

Random forest is an ensemble model build by grouping multiple decision trees. Every single decision tree is trained on a subset of the input data, and the output of different trees are then grouped to produce an output. Random forest often performs very well as a classifier due to its ability to reduce over fitting and handle high-dimensional data [10].

IV. CONVOLUTIONAL NEURAL NETWORKS

While handling image data CNNs are the best type of approach to perform classification. The below mentioned models are used in plant disease detection task.

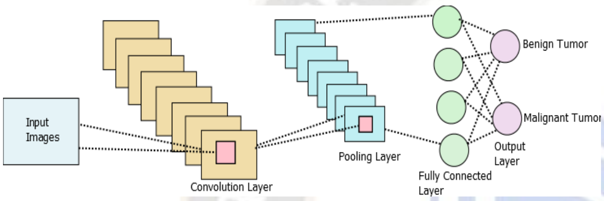

Fig 3: Convolutional Neural Networks

A. VGG16

After making its public debut at the ILSVRC 2014 Conference, the VGG-16 pre-trained model for image classification has swiftly established itself as one of the most popular choices among those that are currently accessible. It was built at the Visual Graphics Group at the University of Oxford, and it quickly became the standard model for picture classification jobs, surpassing the AlexNet that had been utilized up until that point [10].

B. ResNet50

The ResNet50 model has a total of 48 convolution layers, including one Max Pool layer, one Average Pool layer, and one more convolution layer than that. These layers are piled one on top of the other in this structure. The designation "ResNet50" has been best owed upon the version of ResNet that is now in service. The ResNet model in question has been the focus of a considerable amount of research, and its construction involves 3.8 x 109 distinct floating point operations [11]. Computer vision applications such as object identification, object localization, and image categorization all perform satisfactorily when applying this approach. Applications that use computer vision are also known as applications that use computer vision. Figure 6, which provides a graphical depiction of the architecture of ResNet50 and can be viewed at any time, can be found on this website. You can look at it whenever you choose.

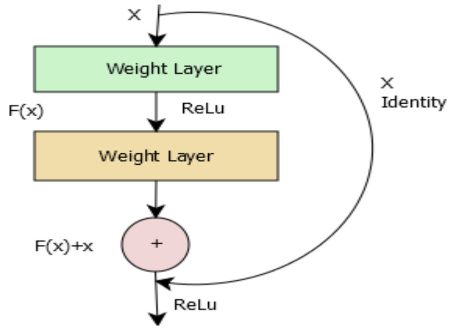

Fig 4: ResNet-50 mode

C. EfficientNetV2

EfficientNetV2 is a type of CNN which scales up depth and width of all dimensions, and uses compound coefficient for resolution. This model needs more layers to increase the receptive field when an input image size is greater. The inverted bottleneck residual blocks are designed to be efficient and effective. They use a smaller number of parameters than traditional residual blocks, while still achieving comparable accuracy. The compound scaling method allows Efficient Net to achieve state-of-the-art accuracy on a variety of image classification tasks. It is a powerful tool for scaling up neural networks without sacrificing accuracy or efficiency [12].

V. PROPOSED CROP RECOMMENDATION SYSTEM

One of the tasks in the categorizing process is to make crop recommendations. In order to classify many different plant species, conventional machine learning (ML) algorithms were utilized. The NPK value of the soil, the temperature of the surrounding environment, the pH of the soil, and the amount of rainfall that occurs in a particular region were among the features that were considered in the process of training the model. The purpose of crop recommendation is to determine the kind of crop that will flourish under a particular set of environmental conditions and advise farmers on how to grow that crop [12].

- The NPK value of the soil is a measurement that can be used to determine how much nitrogen, phosphate, and potassium it contains. The development of plants absolutely requires the presence of these nutrients.
- The temperature of the environment plays an important role in the development of plants. Different plants require varied temperatures.
- The acidity of the soil can be determined by its pH value. Various plant species have preferences for water with varying pH levels.
- The amount of water that each type of plant needs might vary greatly.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8354
***Article Received:** 15 July 2023 **Revised:** 06 September 2023 **Accepted:** 19 September 2023*

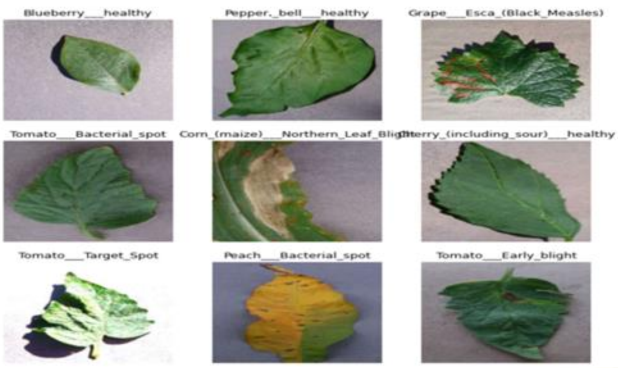

Fig 1: Sample Images of different plant leaves

6. Plant Disease Identification

The dataset that is available for this challenge has 70,000 images in total. Image classification is a demanding task for standard ML systems since feature detection is difficult. Convolutional neural networks (CNNs) are used for difficult image classification applications because they can recognize features during training and provide greater accuracy than traditional ML techniques. We proposed three CNN models, VGG16, ResNet50, and EfficientNetV2 for classification.

VI. RESULTS AND DISCUSSION

The classification accuracy of various ML models are shown in table 1 and CNN models are shown in table 2

Algorithm	Accuracy	Precision	Recall	F1-Score
Logistic Regression	94.54	0.95	0.95	0.94
Decision Tree	97.72	0.98	0.98	0.98
Support Vector Machine (SVM)	9.09	0.59	0.09	0.11
Multi-Layer Perceptron	95.22	0.96	0.95	0.95
Random Forest	99.31	0.99	0.99	0.99

Table1: Accuracy of various ML algorithms

Table 1 concludes that the random forest achieved greater accuracy 99.92%.

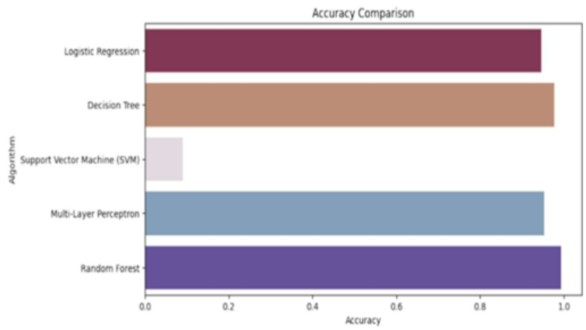

Figure 2: Histogram representation of Crop recommendation system

Figure 3: Accuracy and Loss graphs of the VGG model

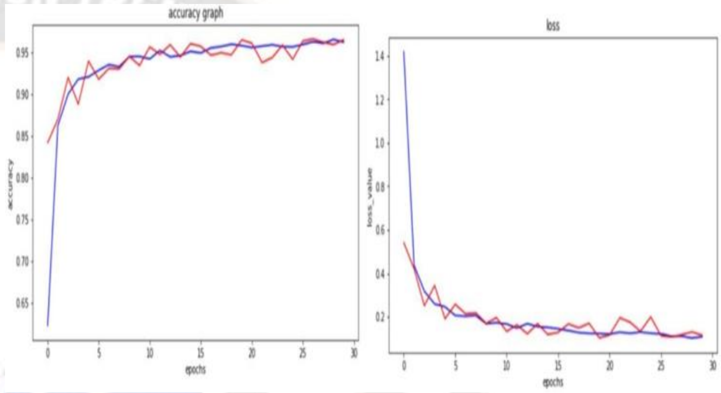

Figure 4: Accuracy and Loss graphs of the ResNet50 model

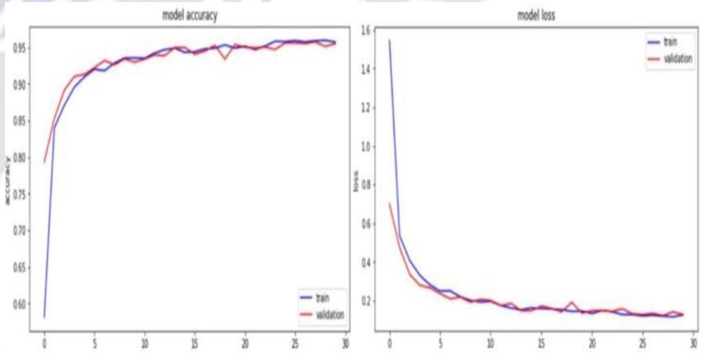

Figure 5: Accuracy and Loss graphs of the EfficientNetV2 model

Table2: Classification accuracy of CNN models

Architecture	Training Accuracy	Validation Accuracy	Testing Accuracy
VGG16	92.19	91.34	91.79
ResNet50	96.03	95.42	95.54
EfficientNetV2	96.07	95.54	95.84

Table 2 concludes that VGG16 achieved an accuracy of 91.78%, ResNet50 achieved an accuracy of 95.53% and EfficientNetV2 achieved better accuracy of 95.83% among all other models.

VII. CONCLUSION

This paper will assist in resolving the enormous harvesting issue facing the agricultural sector. Our comparison of five different algorithms shows that the random forest is an efficient model. The overall accuracy of Random Forest is 99.3%. For classification a comparison of VGG16, ResNet50, and EfficientNetV2 was provided. EfficientNetV2 performed better than VGG16 and ResNet50, with an accuracy of 96.07%. Infuture CNN can be replaced by ViT, CoAtNet models to improve the classification accuracy .The plant leaves must be collected from different regions to improve multi class disease classification.

REFERENCES

[1] Kulkarni, P., Karwande, A., Kolhe, T., Kamble, S., Joshi, A. and Wyawahare, M., 2021. Plant Disease Detection Using Image Processing and Machine Learning.ar XivpreprintarXiv:2106.10698

[2] G. Chauhan and A. Chaudhary, "Crop Recommendation System using Machine Learning Algorithms," 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART), 2021, pp. 109-112, doi: 10.1109/SMART52563.2021.9676210.

[3] Sharada Prasanna Mohanty, David Hughes, Marcel Salathe, 2016, Using Deep Learning for Image-Based Plant Disease Detection.arXiv preprint arXiv:2106.10698.

[4] HassanSM,MajiAK,Jasi´nskiM,LeonowiczZ,Jasi´nskaE. Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach. Electronics. 2021; 10(12):1388

[5] Cadenas, J. M., Garrido, M. C., & Martínez-España, R. (2023). A Methodology Based on Machine Learning and Soft Computing to Design More Sustainable Agriculture Systems. Sensors, 23(6), 3038.

[6] Kaneda, Y., Shibata, S., & Mineno, H. (2017). Multi-modal sliding window-based support vector regression for predicting plant water stress. Knowledge-Based Systems, 134, 135-148.

[7] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. Frontiers in plant science, 7, 1419.

[8] Chen, J., Chen, J., Zhang, D., Sun, Y., & Nanehkaran, Y. A. (2020). Using deep transfer learning for image-based plant disease identification. Computers and Electronics in Agriculture, 173, 105393.

[9] Dhakal, A., & Shakya, S. (2018). Image-based plant disease detection with deep learning. International Journal of Computer Trends and Technology, 61(1), 26-29.

[10] Saleem, M. H., Khanchi, S., Potgieter, J., & Arif, K. M. (2020). Image-based plant disease identification by deep learning meta-architectures. Plants, 9(11), 1451.

[11] Ahmad, N., Asif, H. M. S., Saleem, G., Younus, M. U., Anwar, S., & Anjum, M. R. (2021). Leaf image-based plant disease identification using color and texture features. Wireless Personal Communications, 121(2), 1139-1168.

[12] Sharath, D. M., Kumar, S. A., Rohan, M. G., & Prathap, C. (2019, April). Image based plant disease detection in pomegranate plant for bacterial blight. In 2019 international conference on communication and signal processing (ICCSP) (pp. 0645-0649). IEEE.