

Prediction of Leaf Diseases by using Machine Learning Techniques-A New Approach to Applied Informatics

R. Ravikumar,
Assistant Professor in Information Technology,
Sri Ramakrishna Mission Vidyalaya
College of Arts and Science,
Coimbatore-641 620.Tamilnadu, India.
E-mail id: rravigrat@gmail.com

Dr. V. Arulmozhi,
Associate Professor,
Department of Research-Ph.D
Computer Science, Tirupur Kumaran College for Women,
Tiruppur-641 687, Tamilnadu, India.
E-mail id: drarultkc@gmail.com

Abstract--Applied Informatics is a new emerging field which encompasses information technology and other areas of science. Many Automatic detection of a plant disease is proving their benefits in more fields of plant leaves. Proposed work focus on using machine learning techniques with multilayer Perceptron and simple K-Means algorithm for predicting sugarcane leaf disease by using Weka tool and the obtained results are promising.

Keywords- Machine Learning, Image Processing, Weka Tool, MultilayerPerceptron, SimpleKmeans

I. INTRODUCTION

An Informatics is the information processing structure of natural scientific and engineered systems. It always interconnected the relevant information with the human intervention. The informatics systems developing their theoretical performance in different computer studies atmosphere. By this concept the Applied Informatics inattentive on many speculative approaches with assorted application fields. Applied Informatics spotlight to develop their multipart disciplines of research at extensive features in technical and technological, production and communal fields.

The technology used in this paper is Machine Learning. The Machine Learning is a field from the computer science it processes the computers learning ability by explicit programming skills. This Machine Learning technique was evolved from the Artificial Intelligence concepts. To make some predictions in the data that is in a large set of information this machine learning technique will studies and constructs algorithms. To make decisions and predictions these algorithmic programming instructions overcomes statically by building input models. In this paper the machine learning technique is using Image Processing.

Hence Image Processing is processing the images with different form of mathematical functions and operations by some signal processing. It is usually refers to digital image processing to acquits the images for producing the input images. The digital image processing is used to perform the computational algorithms in digital images. The digital image processing allowing the wider range of algorithms for applying the input data and which can avoid the problems by building up of a noisy signal distortion during the processing. The digital image is a two dimensional image which has a values by pixels and

pictures. So by this image processing similar sets of plant leaf diseases are identify from various plants.

The plant leaf disease is a major problem arising now days in an agricultural field. It produces large number of diseases by some insects affecting their parts. It is very difficult to identify the infected areas from a leaves in a short time span. Also the manual calculations were not in a sufficient manner. Mostly it is difficult to identify the infected parts of a leaf by manually at a short time. To overcome this difficulties a major technical work is to be initiated to complete the work much faster and accurate. The image processing is taking place about this major problem by the help of machine learning tools. And in this paper the tool using here is WEKA tool.

The Weka means Waikato Environment for Knowledge Analysis it is a Machine Learning Software in java language. Weka solves many real world data mining problems. Here the large and small volumes of data from the datasets are recognized. This is a process of machine learning methodology. By using this Weka tool in this paper a Sugarcane plant leaf disease will be identifying very clearly and accurately in algorithmic program identifications.

This manuscript organize as follow: Section II focus review of literature, Section III explain the problem identification in Sugarcane Plant leaf disease, Section IV describes proposed work by using Sugarcane dataset and finally Section V conclude this paper.

II. REVIEW OF LITERATURE

Indian country has a lot of its varieties of trees and plants by their characters and its behaviours. These things mostly specify their parts. So all those parts are having the defects by different diseases which will be affecting in large numbers.

For all these reasons insects is the major problem to spoil everything under the trees and plants. The diseases in the plants are caused by two reasons; the first thing is by Biotic and another one is by Abiotic. The Biotic disease will affect the plants leaf and its stems. And Abiotic disease will occur due to soil compaction, wind and unhealthy soil and root [1].

The Pathogenesis is a beginning stage of a disease which specifies the hosting tissues. In the reproductive stage these infections may be spread from one part of a plant to other parts. The composition of a wet soil and over temperature the plants growing conditions will get poorly [2].

A. Review on Plant Disease:

Prakash M. Mainkar Ed al proposed that, In India Agriculture is an important role for the major development if it is any decreases in agricultural products then the total economical process will be affected. So the essential things such as soil, water and fertilizers are the input resources for sustainability. And here the major growth was affected by some diseases. And by this the farmers are following the chemical test to detecting and classifying plant diseases. In all the developing countries the farming lands are having much but the farmers are to be in less numbers to control the each and every plants in all the days also they are not having the knowledge of non-native diseases. To overcome this problem a consultation with experts are more costly and also the time is consuming. Then the unnecessary usage of pesticides gives danger to natural resources.

The two main categories having in machine-learning techniques it must be the speed and Accuracy. The technology to be needed for an automatic plant disease detection using image processing technology. This technique will give an alert to the farmers in a good time before the disease will be spread [3].

Mrinal Kumar Ed al proposed that, in our India the backbone of the economy is agriculture and which gives impacts to farmers by many diseases. The disease is detected in naked eyes is very difficult and also an inaccurate results will come then the team efforts are in large numbers. To overcome this problem an automatic disease detection system is needed and it is very accurate to give the answers in less time. By this in first we acquire any leaf images in a digital camera and then secondly we pre-processing those images for enhancing the image quality. And by using the classification, segmentation and clustering techniques identifying the pixel and masked it in an infected portion. And finally statistically analyzing the reports [4].

Smita Naikwadi and Niket Amoda proposed that, a plant disease can cause reduction in quality and quantity of agriculture products. Farmers having diversity to select suitable crops. The technological support will be needed for producing the quality. Here the automatic classification of leaf disease is worked out [5].

Mr. N.S. Bharti Ed al, proposed that a plant leaf disease will significantly reducing the quality and quantity of an agriculture products. In approximately 185 million USD was spent for control the plant leaf diseases. An expert

will continuously watching the process is giving much expensive in larger farming areas. Automatic plant leaf disease detection is very important in monitoring large fields. A machine based detection of plant leaf disease will give ideas to control them on early stage. Here they proposed an automatic detection of plant leaf disease using ANNs and k-means algorithms [6].

B. Weka Tool – An Overview:

In this paper by the study of a literature the major defects of a plant disease will be rectified by the different machine learning techniques. So for the studies have been concerned here the technology which is introducing is Weka tool. The description of this tool was already mentioned in an introduction part so with this concept here the review of a literature study regarding this Weka tool will be proceeding now.

A. Rajalakshmi Ed al, proposed that the data pre-processing is the essential step for knowledge discovery. It is to improving the quality of data, the continuous attributes in discretely requiring data mining. It is for the finite set of intervals to generate distinct values. This paper gives continuous values of iris data set using WEKA Tool by various classification algorithms so it shows the discretization improvements in classification accuracy in iris data set [7].

Mangesh Metkari Ed al, proposed that in the medical diagnosis the data mining technique was used. The problem here is to execute the correct disease diagnosis in different symptoms from the patient. Now to classifying the medical datas there are many soft computing methods are used. For this a test should be needed for different diagnosis of diseases. For this data classification is needed to split the dataset by different classes. Here the WEKA Tool is classifying the dataset using some algorithms like Nearest Neighbors Random Forest and J48 which extracts the models to perform the classification on different datasets [8].

Priyanka Sharma proposed that this paper presenting a comparative analysis of different decision tree classification algorithms using WEKA Tool. A Classification process is used and also the various algorithms are compared. Here WEKA provides SQL databases using java Database connectivity and it process the results. The simple table suitable processing for collecting linked database tables. In this the main user interface is an explorer for functionality common based interface [9].

III. PROBLEM IDENTIFICATION

In this paper a Sugarcane Disease problem was discussed. The Sugarcane disease is manually identifying the disease syndrome. Here by an experience a particular man has been examining and identifying the problems which occurring at the sugarcane plant disease manually. So through this process mostly a misguidance of plant leaf disease symptoms will be identified. Then to tell a proper accurate and automated process of leaf disease identification

was accessing through some machine learning techniques here. So by these activities this proposed work provides the automatic system for identify the Sugarcane disease and calculate an evidently proper control results for the Sugarcane disease.

IV. PROPOSED WORK

A. Description of Sugarcane:

Sugarcane is a steamy, perpetual pasture that forms on the side shoot at the bottom to produce manifold stems, naturally three to four m (10 to 13 ft) high and about 5 cm (2 in) in diameter. The stems grow into cane stalk, which when grown-up constitute around 75% of the whole plant. A mature stalk is typically collected of 11–16% fiber, 12–16% soluble sugars, 2–3% nonsugars, and 63–73% water. The average yield of cane stalk is 60–70 tons per hectare (24–28 long ton/acre; 27–31 short ton/acre) per year. However, this figure can vary between 30 and 180 tons per hectare depending on knowledge and crop management approach used in sugarcane cultivation. In this survey a sugarcane dataset is collected and it is described in Table-I and table as follows.

Table-I Description of the Sugarcane Dataset

Dataset	Number of Attributes	Number of Instances
Sugarcane Dataset	6	180

Table-II Sugarcane Dataset Attributes

S.No	Attributes	ID
1	S	S1
2	N	S2
3	R	S3
4	X	S4
5	Var	S5
6	block	S6

Table-III Representation of Attributes for Instances

S.No	Label	Count
1	A	45
2	B	45
3	C	45
4	D	45

B. Experimental Result Analysis using Weka:

Weka is a machine learning algorithm and by using GUI which is directly accessing the dataset. It is mostly performing for classifications, clustering and with

Association rules to develop a new machine learning techniques. Weka files are imported from file formats like ARFF, CSV, C4.5 and binary. In this study using Weka tool to predict diseases of Sugarcane. The prediction is based on Multilayer Perceptron, SimpleKmeans.

i) Multilayer Perceptron:

The Multilayer Perceptron (MLP) which is also called as Feed Forward Artificial Neural Network. This concept sets appropriate outputs. It consists of appropriate nodes by directed graphs; here each node is connected to a node. Each node is an neuron with nonlinear functions. To training a network which uses the back propagation algorithms also it is not linearly separable.

It is referred to as “vanilla” neural networks, hence it has a single hidden layer. In all the neurons a multilayer Perceptron has a linear activation function. A weighted inputs will maps the each neurons output. Any number of layers is proving with the standard algebra by the two-layer input and output model. A nonlinear activation function is develop the frequency of potential actions of biological neurons in the brain will make a multilayer Perceptron in different manner [10].

By using a weka tool here in this proposed work a summary of the resultant datas are shown followed by detailed accuracy and confusable matrices.

Table-IV Summary of the result

Correctly Classified Instances	180	100%
Incorrectly Classified Instances	0	0%
Kappa statistic	1	
Mean absolute error	0.0105	
Root mean squared error	0.0142	
Relative absolute error	2.81%	
Root relative squared error	3.28%	
Coverage of cases (0.95 level)	100%	
Mean rel. region size (0.95 level)	25%	
Total Number of Instances	180	

Table-V Detailed Accuracy

TP Rate	FP Rate	Precision	Recall	F-Measure	MC C	ROC Area	PR C Area	Class
1	0	1	1	1	1	1	1	A
1	0	1	1	1	1	1	1	B
1	0	1	1	1	1	1	1	C
1	0	1	1	1	1	1	1	D
Weighted Avg.	0	1	1	1	1	1	1	1

Table-VI Confusion Matrix

A	B	C	D	Classes
45	0	0	0	a=A
0	45	0	0	b=B
0	0	45	0	c=C
0	0	0	45	d=D

ii) SimpleKmeans:

SimpleKmeans method is a vector quantization which is from signal processing it is popular for an cluster analysis in the data mining concept. A partition of n observation with k clusters as a prototype. The resultant of a Voronoi cells.

It is difficult to identify the problems computationally but there is an efficient algorithm to build in a local optimum. They use cluster centers for model data to compare clusters of spatial extent.


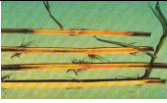
A k-nearest neighbor has loose relationship in popular machine learning techniques for classification which is totally confused with k-means because of the k in the name. The nearest neighbor classifier for a cluster center may be the person can be applying in existing clusters for a new data [11].



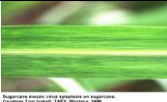

By using a weka tool here in this proposed work a summary of the resultant datas are shown below.

Table-VII Summary of the result

Attribute	Full Data	Cluster
S1	90.5	48.64
S2	118.1444	115.80
S3	20.2556	22.85
S4	11.9389	11.5824
S5	23	17.989
S6	A	A

Table-VIII Sugarcane Disease, Image and Description

S.No	Name of the disease	Images	Description of Syndromes
1	REDROT		prolonged storage or exposure to high relative humidity, environmental pollution, and high temperature
2	WHIP SMUT		multicellular fungi characterized by their large numbers of teliospores

3	WILT		affect the vascular system of plants
4	RINGSPOT		yellow bud mosaic in peaches, yellow vein in grapes, and stunted growth in gladiolus and Narcissus
5	MOSAIC		different genotypes in one individual, who has developed from a single fertilized egg
6	RATOON STUNT		The presence of pin head like orange coloured dots of bacteria on the internal soft tissue in the nodal region

V. CONCLUSION

In this paper the proposed work done by using Weka tool and selected two classifiers namely Multilayer Perceptron and SimpleKmeans algorithms and their results are promising.

ACKNOWLEDGEMENT

I express my sincere thanks to my Supervisor Dr.V.Arulmozhi, Associate Professor in the Department of Research-Ph.D Computer Science, Tirupur Kumaran College for Women, Tirupur, and Tamilnadu, India.

REFERENCES

- [1] Plant Pathology, Web Page Wikipedia
- [2] Malcolm C.Shurtleff, Arthur Kelman, Michael J.Pelczar: PlantPathology,EncyclopediaBritannica <https://www.britannica.com/science/plant-disease>
- [3] Plant Leaf Disease Detection and Classification Using Image Processing Techniques Prakash M. Mainkar1, Shreekanth Ghorpade2, Mayur Adawadkar3Department of Electronics & Telecommunication, MAEER's Maharashtra Institute of Technology, Pune, India 1 prakash.mainkar@mitpune.edu.in, 2 gshrikant41@gmail.com,3mimayur2001@gmail.comInternational Journal of Innovative and Emerging Research in Engineering Volume 2, Issue 4, 2015 ;
- [4] Wheat Leaf Disease Detection Using Image Processing Mrinal Kumar1 , Tathagata Hazra2 , Dr. Sanjaya Shankar Tripathy3 1, 2 ME Student, Department of ECE, BIT Mesra, Ranchi 3Assistant Professor Department of ECE BIT Mesra, Ranchi; International Journal of Latest Technology in Engineering, Management & Applied Science (IJLTEMAS) Volume VI, Issue IV, April 2017 | ISSN 2278-2540.
- [5] ADVANCES IN IMAGE PROCESSING FOR DETECTION OF PLANT DISEASES Smita Naikwadi, Niket Amoda Department of Electronics and Telecommunication, K.C. College of Engineering,

- Management studies and Research, University of Mumbai, India; International Journal of Application or Innovation in Engineering & Management (IJAIEM) Web Site: www.ijaiem.org Email: editor@ijaiem.org, editorijaiem@gmail.com Volume 2, Issue 11, November 2013 ISSN 2319 – 4847
- [6] Detection and Classification of Plant Diseases Mr. N.S. Bharti, Prof. R.M. Mulajkar. 1 ME Student, Electronics and Telecommunication, J.C.O.E. Kuran, Maharashtra, India 2Professor, Electronics and Telecommunication, J.C.O.E. Kuran, Maharashtra, India, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 02 Issue: 02 | May-2015 www.irjet.net p-ISSN: 2395-0072
- [7] Data Discretization Technique Using WEKA Tool A. Rajalakshmi¹, R. Vinodhini², K. Fathima Bibi³M.Phil Research Scholar^{1, 2}, Assistant professor³Department of Computer Science Rajah Serfoji Govt. College (Autonomous), Thanjavur-5. A. Rajalakshmi et al | IJCSET (www.ijcset.net) | August 2016 | Vol 6, Issue 8, 293-298.
- [8] Improve the Classification Accuracy of the Heart Disease Data Using Discretization Mangesh Metkari* Madhavi Pradhan AISSMS Coe Pune, Pune University AISSMS Coe Pune, Pune University, International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163 Issue 10, Volume 2 (October 2015).
- [9] [9] Priyanka Sharma, “Comparative Analysis of Various Decision Tree Classification Algorithms using WEKA”, February 15 Volume 3 Issue 2 , International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), ISSN: 2321-8169, PP: 684 - 690, DOI: 10.17762/ijritcc2321-8169.150254 APA.
- [10] https://en.wikipedia.org/wiki/Multilayer_perceptron
- [11] https://en.wikipedia.org/wiki/K-means_clustering