Leaf Disease Classification using Machine Learning Approach

R. B. Dhumale[#1], N. D. Thombare[2], Mr. Amit M Umbrajkaar[3], A. N. Sarwade[4]
1, 2, 4 Department of Electronics and Telecommunication Engineering, Sinhgad College of Engineering, Pune.
3Department of Mechanical Engineering, Dr. D. Y. Patil Institute of Engineering, Management & Research, Akurdi, Pune.

Abstract— **Well-timed and precise leaf diseases classification of plant is the critical factors in agriculture. This leads to harm the crops and there may be financial losses in agriculture. Due to lack of knowledge in the leaf diseases of plant, farmers are using extreme pesticides. Plant diseases are microscopic in nature hence the diagnosis of plant disease is controlled by visual capabilities of human being. The visual nature of leaf disease classification emphasizes the computer visualization. This paper proposes a comparative study of leaf disease classification using Image Processing and Machine Learning Techniques. This approach of disease classification consists is implemented using RGB value Extraction and Disease Classification. The two diseases are classified viz. Whiteflies and Rust. The training database of leaf images consist input as RGB values of pixels and target values which indicates diseases of leaf like Healthy, Whiteflies and Rust. The different learning algorithms are discussed, trained and tested over 310 images. Comparison of disease classification is given for different classification techniques.**

Keyword- Leaf disease, rust, whiteflies, RGB model, artificial neural network, learning rules.

1. INTRODUCTION

Leaf disease has become a most important threat in worldwide food safety. Plant diseases add **10–16% losses** in the worldwide harvest of crops and every year loss is predictable **US$220 billion** [1]. Usually, plant diseases harms the photosynthetic apparatus and have an effect on the growth of plant [2]. The majority (85%) of plant diseases are caused by **fungal or fungal-like organisms**. Additional severe diseases of plants are originated by **bacteria, viruses, and viroids** and a small amount of diseases are caused by some **nematodes** [3]. **Spectral signatures** of a leaf could not be investigated properly using parametric approaches **like simple or multiple regression and functional statistics**. As a result, non-parametric methods such as **Principal Component Analysis (PCA), Fuzzy Logic (FL), Support Vector Machine (SVM), Cluster Analysis (CA), Partial Least- Square (PLS), and Neural Networks (NNs)** have been adopted in the area of image processing for leaf disease classification [4]. PCA is a multivariate statistical technique that reduces redundancy in univariate study. PCA aids to recognize patterns of spectral data. Mainly, PCA converts huge numbers of correlated variables into lesser number of uncorrelated variables [5]. The PCA and **Partial Least-Square (PLS)** are recently highlighted for recognition of fungal diseases of wheat and barley [6].

Whetton et al. Proposed use of PCA on healthy and yellow rust and fusarium infected cereal crops at various phases of growth and learned their temporal pattern and serial autocorrelation. PLS is capable for correct prediction to every growing phase of plant [7]. In the second part it is concluded that the regression model implemented for fusarium head blight and yellow rust in wheat can be used to forecast these diseases in barley.

The method proposed in [8], is based on **Image Processing (IP)** approach by using spread on leaves by using FL without human intervention rating. The results are proved the precise and acceptability as compared to the manual grading.

An easy and computationally capable technique used for leaf disease identification and grading using machine vision approach is proposed in [9]. This method is divided into plant identification using **IP-ANN** and classification of diseases present in the leaf using **K-Means based segmentation** of defected area. The grading of disease is done on the basis of the quantity of disease present in the leaf.

The combination of Back-Propagation Neural Network (BPNN) with a Self-Organizing Feature Map (SOM) is utilized to identify colors of grape leaf [10]. This system has capability to categorize the grape leaf diseases as **Scab, Rust and no disease**. Still there are several limitations of extracting vague color pixels from the background of the image. The system shows potential performance for several agricultural product analyses.

Rice brown spot leaf disease is identified using BPNN in [11]. The images of rice leaves are obtained from the northern part of Ningxia Hui. The color features of diseases and healthy region are given as input to BPNN. The obtained result shows the capability of this method. The robust detection & classification method for fast and correct is proposed in [12]. Initially RGB images of leaves obtained then a color transformation structure is formed for the RGB-image. Then image segmentation, boundary identification and classification using ANN is carried out.

From above survey it is observed that different techniques are used to classify the leaf diseases using ANN. An ANN functions like the human brain with the capability of learning, reasoning and self-correction [13]. Learning is the main task in ANN to get more accurate results. The learning rules have different outcomes, different mathematical model and applicability; these rules are shown in TABLE I [14]. There are six learning rules as given below [15]:

- **Hebbian Learning Rule** – It recognizes how to change the weights of nodes in ANN.
- **Delta Learning Rule** – Change in synaptic weight of a node is identical to the product of error and the input.
- **Competitive Learning Rule** – The relationship rule is the supervised learning.
- **Memory based Learning** – all the previous experiences are store in large memory.
- **Outstar Learning Rule** – It is used in such cases where it is assumed that nodes or neurons in an ANN organized in a layer.
- **Boltzmann Learning Rule** – The learning is a stochastic based.

In this paper ***Leaf Disease Classification using Delta Learning Rule (LDCDLR)*** is proposed. The RGB values of pixels are input and three conditions like healthy, rust and whiteflies are the output. The Delta Learning Rule is used to train ANN. The results of Delta Learning Rule are compared with the Maximum likelihood classifier.

TABLE I
SUMMARY OF LEARNING RULES [14]

Learning Rule	Weight adjustment	Initial Weight	Learning*
Hebbian	$\Delta W_{Kj}(P) = \eta X_j(P)\ Y_K(P)$	Non Zero	U
Delta	$\Delta W_{Kj}(p) = \eta\ E_k(p) X_j(p)$	Random	S
Competitive Or Winner-Takes-All Or Instar	$\Delta W_{jk} = \eta\ (a_j - w_{kj})$	Random but Normalized	U
Outstar	$\Delta W_{Kj} = \begin{cases} \eta\ (Y_K - W_{jk}) & j \text{ wins} \\ 0 & j \text{ losses} \end{cases}$	Zero	S
Boltzmann	$\Delta W_{Kj} = \eta(\overline{P_{kj}} - P_{jk})$	Random	S
*U= Unsupervised, S= Supervised.			

2. IMAGE CLASSIFICATION

Image is composed of multiple bands in Electromagnetic (EM) spectrum. All bands together form image which is spectral information. ***Image classification is converting this spectral information into particular information class*** [16]. ***Spectral signature*** of each class is different. The difference in the ***Near to Infra Red (NIR)*** images' intensities is not just due to the particular color of the material, but also absorption and reflectance characteristics of the colorant. This relative independency of NIR and color information makes NIR images a prime candidate for classification [17]. ***Image classification*** is also defined as assigning each pixel in the image to categories or classes of interest [18]. Examples: built-up areas, water-body, green vegetation, bare soil, rocky areas, cloud, shadow, etc., in false color composite composition of three bands of satellite data. Following are the methods of image classification [19]:

- Visual Interpretation.

- Digital Image Classification.
- Hybrid Process.

In order to classify a set of data into different classes or categories, the relationship between the data and the classes into which they are classified must be well understand. To achieve this by computer, the computer must be trained. ***Training is the key to success of classification*** [20]. Classification techniques are originally developed for research in Pattern Recognition field [21]. Important aspects in pattern recognition of accurate classification are ***learning techniques and feature sets*** [22]. The learning techniques are ***supervised learning*** and ***unsupervised learning***.

2.1 Supervised Learning:

Learning process designed to form a mapping from one set of variables (data) to another set of variables (information classes). A teacher is involved in this learning process. Following are methods of supervised learning.

- Minimum distance to means classification algorithm.
- Parallelepiped classifier.
- Maximum likelihood classifier.

2.2 Unsupervised Learning:

This learning happens without a teacher. This method is exploration of the data space to discover the scientific laws underlying the data distribution.

2.1.1 Minimum Distance to Means Classification (MDMC):

In ***MDMC***, it is observed that the ***test sample*** is closest to which input vector. Where train vectors are classified into some groups as shown in Fig. 1. Some distance measures are required to find out closest distance; simplest distance measure is *"Euclidian Distance"*. All together these groupings of pixels are done on the basis of particular observations. Sometimes these observations may be false. As shown in Fig. 2, all these responses in Group-1 are shown by red color but one entry is with blue color which is wrong observation in Group-1. Such entry is called as ***"Outlier"*** in pattern presentation process. If *"Nearest Neighbor Criteria"* is adopted then nearest neighbor will be Outlier**.** So for avoiding such observations it requires to define in pattern presentation process as shown in Fig. 3 which is known as **"Local Neighborhood"**. Average of all the samples in one group is considered and distance of test sample is measured from this average. This method is called as **Minimum Distance to Means Classification.**

This method normally classifies every pixel no matter how far it is from a class mean unless the minimum distance condition is applied.

2.1.2 Parallelepiped Classifier (PPC):

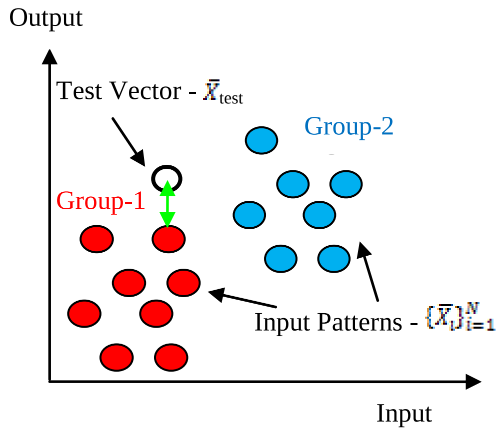

Fig.1. Concept of Minimum distance to means classification

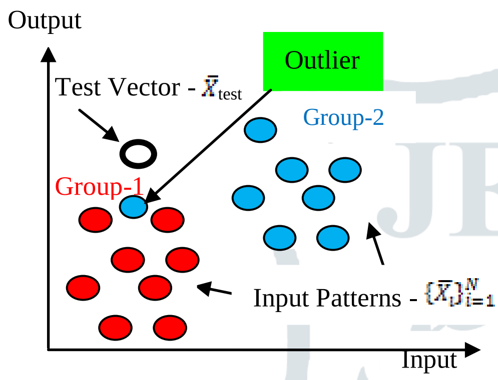

Fig.2. Concept of Outlier in MDMC

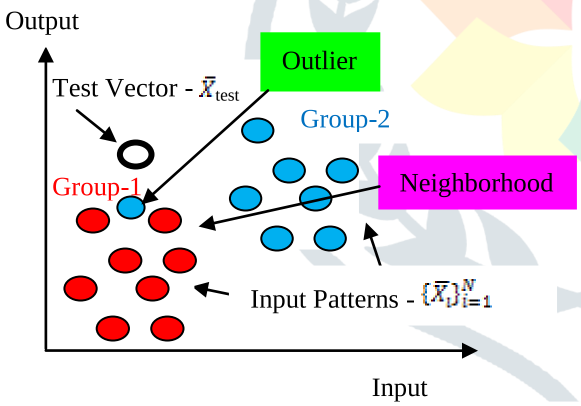

Fig.3. Concept of Neighborhood in MDMC

In the ***Parallelepiped classifier (PPC)*** instead of considering just mean distance value, produce a box in the region of the pixels related with particular group by taking into account ***minimum and maximum variance*** of the pixels for class itself. In the PPC minimum and maximum variance is considered, hence this PPC is really a ***"feature space"*** classifier. A pixel is assigned to a class only if its feature vector is matching with the any one group.

2.1.1 Maximum Likelihood Classifier (MLC):

In ***Maximum Likelihood Classifier (MLC)***, let us consider there are *C* classes denoted by ω_j and in each class different numbers of samples. X is the feature vector. The *Class Conditional Density* $p(X/\omega_j)$ is calculated which is represented by Gaussian Distribution Function given by Eq. 1.

$$p\left(\frac{X}{\omega_j}\right) = N\left(\mu_j, \Sigma_k\right) \quad (1)$$

Class Conditional Density is calculated using Eq. 2.

$$p\left(\frac{X}{\omega_j}\right) = \frac{1}{\delta_j\sqrt{2\pi}} e^{\frac{1}{2}(\frac{X-\mu_j}{\delta_j})^2} \qquad (2)$$

Where μ_j and δ_j are maximum value and distance as shown in Fig. 4.

3. Leaf disease classification using delta learning rule

The proposed **Leaf Disease Classification using Delta Learning Rule (*LDCDLR)*** method is shown in Fig. 5 which consists of following steps:

- Select pixel to convert into RGB,
- Read RGB value of Pixel,
- Classify using Feed Forward Back Propagation Neural Network (FFBPNN).

3.1 Select pixel to convert into RGB:

The leaf image is input to the *LDCDLR*. The pixel is selected from leaf to classify into three categories like **Healthy, Rust and Whiteflies**. The samples of images for these diseases are shown in Fig. 6.

3.2 Read RGB value of Pixel:

The RGB value of selected pixel takes from the current axes selected as shown in Fig. 7. The pixel can be selected using a mouse. Pixels are entered by pressing a mouse button or any key on the keyboard excluding carriage return, which stops the input before pixel are entered. The pixel values are collected for diseases like **Whiteflies, Rust** including **Healthy** condition. Some values are shown in Table II such 310 images and 930 pixel values of RGB-Images are collected.

3.3 Disease Classification:

ANNs are trained using above mentioned data learning rules to carry out significant tasks like grouping, recognition or relationship. The working of Artificial Neural Network (ANN) is very closely analogous to the working of a ***Biological Neural Networks (BNNs)***. The *Biological Neural Network* contains the trillions of neuro cells and billions of interconnections between them. Neuro cells work as a "*processing unit*". A single biological neuro cell is shown in Fig. 8. It is connected to many other neuro cells with transmission lines called "*Axons*".

These neuro cells interact with other neuro cells with the help of *synaptic terminals*. Input terminals to the neuro cells are called as *synaptic inputs*. Similarly, the output of the neuro cells is called as synaptic output. Biological neural networks form a tree like structure which consist a long chain of neuro cell working as *nodes*. All these neuro cells work parallelly hence it is called as "*Massively parallel structure*" or "*Extremely parallel structure".* An axon of the biological neural cell carries electrical signals. It has high electric characteristics such *as high Resistive and Capacitive value*. The strength of the electrical signal which is being passed through these axons depends upon the strength of the connection. All these connections form "*Pyramidal Structure*".

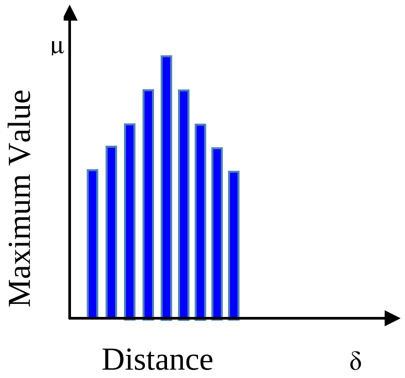

Fig. 4. Gaussian distribution function

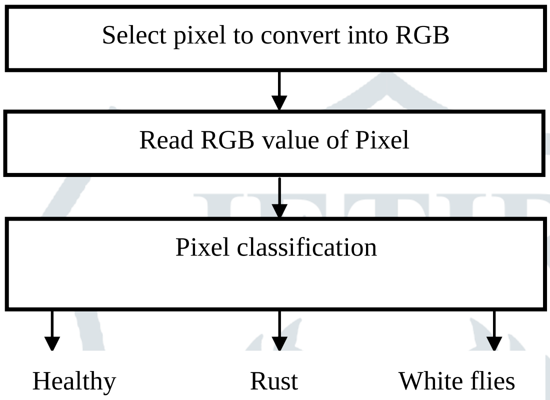

Fig. 5. Proposed methodology of Leaf Disease Classification using Learning Algorithm

a. Rust b. Whiteflies

Fig. 6. Images of Leaf Diseases

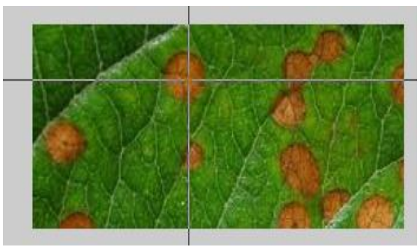

Fig. 7. Selection of pixel to be classified.

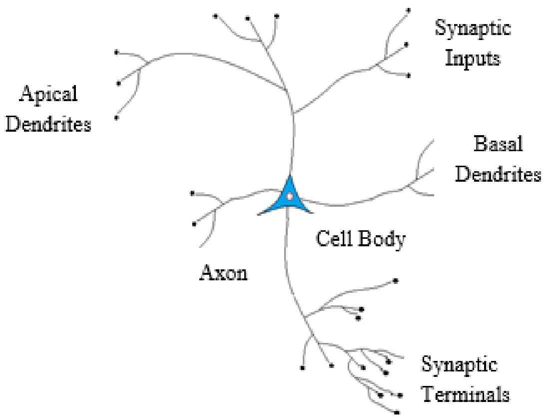

Fig. 8. Biological neural network.

Table II
RGB value for Healthy and Diseases

RGB Values			Condition
R	G	B	
50	108	3	Healthy
55	110	1	
58	116	15	
21	56	5	
32	78	3	
37	86	10	
25	60	1	
179	162	136	Rust
199	184	161	
118	113	51	
137	109	70	
41	41	32	
19	18	12	
12	12	10	
24	16	11	
14	11	8	
114	115	100	Whiteflies
135	134	121	
148	146	136	
158	161	153	
126	126	116	
173	176	159	

Consider an example of teaching English alphabets to a child. If child is taught a pattern of ‘A’, child remembers it similarly if child is taught a pattern of ‘B’ child remembers it and so on. This is exactly like training a neuro cell to store something. The biological neural network learns with the help of feedback and corrects itself after each false output. The storing information is like changing and tuning the electrical signals according to the alphabets being taught. Here for training a biological neuro cell, these electrical signals are the free parameters.

The ANN can be understood using the similar analogy with BNN. Consider the neural network shown in Fig. 9 with ‘n’ number of inputs namely x_1, x_2, x_3, x_4 … x_n. These inputs can be vector. Apart from these n inputs to the neural cell, there is an extra bias input ‘x_0’ given to the neuro cell. This bias input acts as a reference signal for all the ‘n’ inputs. These inputs along with the bias input ‘x_0’ are given to the K^{th} processing unit which in this case is “*summation*”. The free parameters in this case are the weights of the individual transmission lines connecting the neural network to the neuro cell. Weight of connection x_1 to the K^{th} neural cell is shown as W_{k1} and so on till W_{kn}.

The output of K^{th} neuro cell is V_k. V_k is the output signal with some electrical value as given in Eq. 3. V_k is sum value calculated using summation. Hence it is required to convert this value into actual decision. This conversion is carried out using "*Activation function*". Activation function is a function of V_k which gives the output Y_k. Y_k is actual decision of the single neuro cell.

$$V_k = \sum_{j=1}^{n} (X_j W_{Kj}) + (X_0 W_{K0}) \qquad (3)$$

The activation function is the actual decision-making function from which the decision is derived. There are three types of Activation Functions as given below:

- Threshold activation function,
- Linear activation function,
- Sigmoid activation function.

Let us consider some data with input and output. The ANN can be trained using the different learning rules and during learning the difference between input and targeted output is adjusted. This difference in expected value of target point 'T_p' with the obtained point 'Y_p' is termed as error 'E_p' and given by Eq. 4. Similarly, if the errors of multiple points are calculated then the *Mean Square Error (MSE)* is give by Eq. 5.

$$E_p = (T_p - Y_p) \qquad (4)$$

$$E_T = \sum_{P=1}^{q} (T_p - Y_p)^2 \qquad (5)$$

Where q is the size of data. The target of learning rule is to minimize MSE using free parameters. The W_{k0} to W_{kn} are the free parameters which are to be adjusted using learning mechanism. The Delta Learning rule helps to adjust weights using Eq. 6 and Eq. 7.

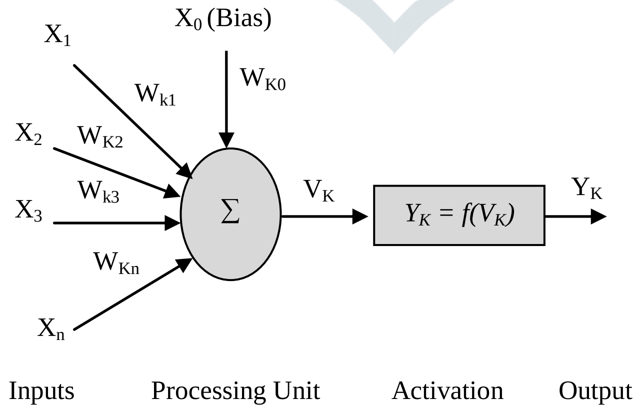

Fig. 9. Architecture of Artificial Neural Network

$$W_{K1} = W_{k1} + \eta \ \frac{1}{n}\sum_{p=0}^{n}(Y_W X_p - Y_p).X_p \qquad (6)$$

$$W_{K0} = W_{k0} + \eta \ \frac{1}{n}\sum_{p=0}^{n}(Y_W X_p - Y_p) \qquad (7)$$

$\frac{1}{n}$ is for creating the derivation procedure a little simple. Where, "η" is learning rate i.e. the rate of fall down or step size of adjusting weight for every iteration. The least MSE point is called as *Local Minima*. This can be achieved by sliding down the initial point to the local minima. A gradual and slow descend leads to slow learning rate and a fast descends leads to fast learning rate. *Fast learning rate may not give the least MSE always while a slow learning rate gives the least possible MSE value but takes a lot of iterations as compared to fast learning rate*. If the value of MSE starts rising due to fast learning rate, it is called as "*Steepest Descend*" and if it goes on decreasing iterations by iterations, it is called as "*Gradient Descent*". So, for finding the least MSE, gradient descent is desirable. In this algorithm, execute much iteration to minimize the MSE value. The maximum limit of iterations is also put to reduce the time and limit the number of iteration executions.

4. Results and Discussion

One of the common problems of supervised learning is the collection of representative training data set for every group. A large amount of time and expenditure can be saved if less number of samples is required for the training of classifier. An added advantage is the potential to decrease training time of classifier. Training data sets are created and given to the MLC and LDCDLR. The comparison of the classifiers for different training data set is given in Table III and Table IV.

It is observed that the accuracy of MLC and LDCDLR is highest for 930 numbers of training samples. If numbers of samples are reduced by 2/3 then the accuracy is 72.29 % and 78.29 % for MLC and LDCDLR respectively. Similarly, if numbers of samples are reduced by 1/5 then the accuracy is 62.29 % and 70.23 % for MLC and LDCDLR respectively.

Table III

Results of the training and testing accuracy

Number of Samples	**Maximum Likelihood Classifier**		**Delta Learning Rule**	
	Training	Testing	Training	Testing
930	99%	89.25%	99.10%	92.92%
750	100%	81.67%	99.98%	85.67%
630	99%	72.29%	97.29%	78.29%
330	100%	67.76%	99.92%	74.76%
186	100%	62.29%	98.24%	70.23%

Table IV

Training and testing time (sec)

Number of Samples	**Maximum Likelihood Classifier**	**Delta Learning Rule**
930	590	8962
750	590	7291
630	590	6529
330	590	3420
186	590	2625

As both classifiers need a sufficient number of samples in each category to illustrate decision boundaries in the feature space, the reduced accuracy is estimated. It might appear perceptive; still, that the LDCDLR would need extra samples than MLC as it suppose no statistical allocation and thus would need extra information to describe these decision regions.

5. Conclusion

Images classification is one of the exigent tasks in analysis of image. Image classification techniques are involved in different areas like remote sensing, medical diagnosis, robotics, etc. Image classification is to identify identical clusters of data points in a specified dataset and conveying it to a class. Different techniques are used to classify images using machine learning approach. In this paper the image classification methods using different learning rules are implemented and tested over 310 images. The images are collected for diseases like *Whiteflies*, *Rust and Healthy leaf*. The comparison of Maximum Likelihood Classifier and Delta Learning Rule is given with respect to accuracy and time. It is concluded that the Maximum Likelihood Classifier requires same time for training and testing of data set with different size but the Delta Learning Rule increases training and testing time with respect to data size. Also the accuracy in both cases increases with increase in data size. The 89.25% and 92.92% accuracy is reported in Maximum Likelihood Classifier and Delta Learning Rule respectively.

6. References

[1] KamleshGolhania, Siva K. Balasundrama, Ganesan Vadamalaib, BiswajeetPradhanc, *A review of neural networks in plant disease detection using hyperspectral data, Information Processing in Agriculture Volume 5, Issue 3, September 2018, Pages 354-371.*

[2] Petit, Anne-Noëlle & Florence, Fontaine & Vatsa, Parul & Clément, Christophe & Vaillant-Gaveau, *Fungicide impacts on photosynthesis in crop plants*. Photosynthesis research. 111. 315-26, Nathalie. (2012).

[3] Tom Howley, Michael G. Madden, Marie-Louise O'Connell and Alan G. Ryder, *The Effect of Principal Component Analysis on Machine Learning Accuracy with High Dimensional Spectral Data*, pp 1-15.

[4] Michael,.MaddenaMarie-LouiseO'ConnellbAlan G.Ryderb, *The effect of principal component analysis on machine learning accuracy with high-dimensional spectral data,* Knowledge-Based Systems, Volume 19, Issue 5, September 2006, Pages 363-370.

[5] Isabelle Guyon, Andre Elisseeff, *An Introduction to Variable and Feature Selection*, Journal of Machine Learning Research, Vol, 3, 2003, pp 1157-1182.

[6] Simone Maccaferri, Annett Klinder , Stefano Cacciatore , Roberto Chitarrari , Harue Honda , Claudio Luchinat , Ivano Bertini , Paola Carnevali, Glenn R. Gibson , Patrizia Brigidi , Adele Costabile, *In vitro fermentation of potential prebiotic flours from natural sources: Impact on the human colonic microbiota and metabolome,* Molecular Nutrition Food Research, 02 July 2012.

[7] Zolo Kiala, John Odindi, Onisimo Mutanga, *Potential of interval partial least square regression in estimating leaf area index*, South African Journal of Science, Volume 113, pp 9-10 September-October 2017.

[8] Sanjeev S Sannakki, Vijay S Rajpurohit, V B Nargund , Arun Kumar R , Prema S Yallur, '*Leaf Disease Grading by Machine Vision and Fuzzy Logic*', Int. J. Comp. Tech. Appl., Vol. 2, Iss. 5, pp- 1709-1716, Oct-2011.

[9] Jagan Bihari Padhy, Devarsiti Dillip Kumar, Ladi Manish, Lavanya Choudhry, *Leaf Disease Detection Using K-Means Clustering And Fuzzy Logic Classifier*, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 6, Iss. 3, pp 1-7, March 2017.

[10] Savita N. Ghaiwat, Parul Arora, *Detection and Classification of Plant Leaf Diseases Using Image processing Techniques: A Review*, International Journal of Recent Advances in Engineering & Technology (IJRAET), Volume-2, Issue - 3, pp 1-6, 2014.

[11] Sannakki S.S., Rajpurohit V.S., Nargund V.B., Arun Kumar R.and Yallur P.S., *A Hybrid Intelligent System for Automated Pomgranate Disease detection and Grading*, International Journal of Machine Intelligence, Volume 3, Issue 2, 2011, pp-36-44.

[12] Haiying Yua,b, Eike Luedelingc , and Jianchu Xu, *Winter and spring warming result in delayed spring phenology on the Tibetan Plateau*, PNAS, December 21, 2010, vol. 107 , no. 51-57,

[13] R. B. Dhumale, M. P. Ghatule, N. D. Thombare, P. M. Bangare, *An Overview of Artificial Neural Networks: Part 1*, CiiT International Journal of Artificial Intelligent Systems and Machine Learning

[14] R. B. Dhumale, *An Overview of Artificial Neural Networks: Part 4 Learning Mechanism*, CiiT International Journal of Artificial Intelligent Systems and Machine Learning. Vol. 10. Iss. 4. March 2018, pp 77-82.

[15] R. B. Dhumale, *An Overview of Artificial Neural Networks: Part 3 Activation Functions*, CiiT

International Journal of Artificial Intelligent Systems and Machine Learning, Vol. 10. Iss. 3. April 2018, pp 66-71.
[16] H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik and Z. ALRahamneh, Fast and Accurate Detection and Classification of Plant Diseases , International Journal of Computer Applications, Volume 17– No.1, March 2011, pp 31-38.
[17] Ilia Parshakov, Automatic Class Labelling of Classified Imagery using A Hyperspectral Library, Department of Geography University of Lethbridge Lethbridge, Alberta, Canada, 2012.
[18] Urška Kanjir, Harm Greidanus, Krištof Oštir, ‚Vessel detection and classification from spaceborne optical images: A literature survey, Remote Sensing of Environment, Vol. 207, 15 March 2018, Pages 1-26.
[19] Siamak Khorram, Cynthia F. van der Wiele, Frank H. Koch, Stacy A. C. Nelson, Matthew D. Potts, *Principles of Applied Remote Sensing*, Springer, Cham, 2016.
[20] Yihua Chen, *Similarity-based Classification: Concepts and Algorithms*, Journal of Machine Learning Research 10, 2009, pp 747-776.
[21] Priyanka Sharm, Manavjeet Kaur, *Classification in Pattern Recognition: A Review*, International Journal of Advanced Research in Computer Science and Software Engineering Research, Volume 3, Issue 4, April 2013, pp 298-306.
[22] M.Subba Rao, *Comparative Analysis of Pattern Recognition Methods: An Overview*, Indian Journal of Computer Science and Engineering (IJCSE), Vol. 2, Iss. 3 Jun-Jul 2011, pp 385-390.