ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 2, February 2018

Extension of SVM for Multi-Class Classification

[1] Rupali A. Mangrule, [2] Dr. R. A. Khan
[1] Asst. Prof, MIT, Aurangabad, [2] Asso. Prof, DIEMS, Aurangabad.

Abstract: **The support vector machine (SVM) is a well-known algorithm for binary classification problems. SVM was originally developed to classify only two classes. In the real world, most of the applications involve multiclass classification e.g. plant disease detection, face detection, plant classification etc. Binary SVM can be extended in different ways for multi-class classification problems. This paper provides a survey of different multi-class approaches like One vs. one, one vs. rest, Directed Acyclic Graph (DAG) and Error Corrected Output Coding (ECOC). The paper also presents how Multi SVM is effectively used to early detect and classify the plant diseases based on plant leaf symptoms. The early detection of plant disease ultimately effects on crop production which is the major factor of the Indian economy.**

Keywords: **- Support Vector Machine, Plant Disease Detection, Image Processing, Classification.**

I. INTRODUCTION

Support Vector Machine (SVM) is well known algorithm for binary classification. Due to the class separation process, and the facilities of kernel space makes SVM a robust and powerful tool for the classification in number of applications. It is a supervised learning technique. For a data set consisting of features set and label set, SVM classifier builds a model to predict classes for new samples.

There are two types of SVM Classifier:

A) Linear SVM classifier

B) Non-linear SVM classifier

In linear classifier, it predicts a straight hyper plane dividing two classes. The objective while drawing hyper plane is on maximizing the distance from hyper plane to the nearest data points of either class. In the real world, dataset is generally dispersed up to some extent. For this Vapnik suggested creating nonlinear classifiers by applying the kernel trick to maximum-margin hyper planes. Figure 1 shows that there are a lot of linear hyper planes which can separate the samples, but there is only one hyper plane that makes the margin largest between the two kinds of sample points.

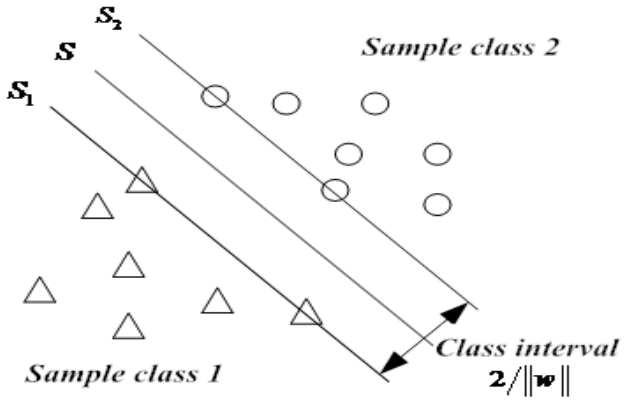

Figure 1

II. SVM FOR MULTI CLASS CLASSIFICATION

Originally SVM model is a binary classifier, it can be extended to solve multi-class classification problems. As the applications of binary classification are very limited, different methods have been proposed where we construct a multi-class classifier by combining several binary classifiers.

A) One-Against-All (OAA)

It constructs k SVM models where k is the number of classes. The M^{Th} SVM is trained with all of the examples in the M^{Th} class with positive labels, and all other examples with negative labels. Suppose the dataset is to be classified into M classes. Therefore, M binary SVM classifiers may be created where each classifier is trained to distinguish one class from the remaining M-1 classes. For example, class one binary classifier is designed to discriminate between class one data vectors and the data vectors of the remaining classes. This method is also called winner-take-all classification.

B) One-Against-One (OAO)

It constructs k (k-1)/2 classifiers where each one is trained on data from two classes. Hence it is also known as 'pair wise classification'. A voting strategy is used after all the k(k-1)/2 classifiers are trained to make a final decision, and this strategy is called max-win strategy .

C) Error Correcting Output Code (ECOC)

The concept of Error- Correcting Output Coding (ECOC) based multi-class method is to apply binary (two-class) classifiers to solve the multi-class classification problems. This approach works by converting M class classification problem into a large number L of 2-class classification problems. ECOC assigns a unique code word to a class instead of assigning each class a label. A (L, M, d) error correcting code is a L bit long, having C unique code words with a Hamming distance of d. The hamming distance between two code words is the number of bit positions in which both

ISSN (Online) 2456 -1304

International Journal of Science, Engineering and Management (IJSEM)
Vol 3, Issue 2, February 2018

differs. In a classification problem M is the number of classes and L is a number decided by the method used to generate error-correcting codes. Several methods such as Hadamard Matrix codes, BCH codes, random codes and exhaustive codes are proposed to generate error correcting codes.

D) Directed Acyclic Graph Support Vector Machines (DAGSVM)

The training phase of this algorithm is the same as the one against- one method by solving k(k - 1)/2 binary SVMs. And in the testing phase, it uses a rooted binary directed acyclic graph which has k(k - 1)/2 internal nodes and k leaves. Each node is a binary SVM of ith and jth classes. Given a test sample x, starting at the root node, the binary decision function is evaluated. Then it moves to either left or right depending on the output value. Therefore, we go through a path before reaching a leaf node which indicates the predicted class.

III. MULTI CLASS SVM FOR PLANT DISEASE DETECTION AND CLASSIFICATION

In India, the main source of income is from agriculture. Due to this reason, the disease detection in plants plays an important role in agriculture field. If disease is detected earlier a proper action can be taken. The existing methods for plant disease detection system are naked eye observation by domain experts. It needs a large team of experts and even they cost high for large farms. Also the experts are not available on time at remote location. Hence it needs automatic plant disease detection and classification simply by seeing the symptoms on the plants leaves. This automatic disease detection system will take less time, less efforts and even very economical in detecting and classifying plant disease based on leaf symptoms.

Methodology: The methodology of proposed work is shown in the block diagram and contains five stages.

Image Acquisition

Devices like digital camera are used to capture images of leafs of different types. Both diseased leaf images and fewer healthy leaf images are taken in order to train the system and also to test the system.

Preprocessing of Image

Images that are taken contain noise. Hence Image preprocessing is done to remove the noise. This step is important because it improves the effectiveness of subsequent tasks like feature extraction.

Image Segmentation

During image segmentation, the given image is separated into a homogeneous region based on certain features. Larger data sets are put together into clusters of smaller and similar data sets using clustering method

Classification

The last step in the proposed method is classification of plant diseases based on leaf samples. The multi class SVM classifier is used to identify classes. Support Vector Machine (SVM) classifier is one of the dominant classification algorithms and gives better accuracy than other classification algorithms.

Figure block Diagram for Proposed System

The binary classifier which makes use of the hyper-plane which is also called as the decision boundary between two of the classes is called as Support Vector machine (SVM). Some of the problems of pattern recognition like texture classification make use of SVM. Mapping of nonlinear input data to the linear data provides good classification in high dimensional space in SVM. The marginal distance is maximized between different classes by SVM. Different kernels are used to divide the classes. SVM is basically binary classifier which determines the hyper plane in dividing two classes. The boundary is maximized between the hyper plane and the two classes. The samples that are nearest to the margin will be selected in determining the hyper plane are called as support vectors.The last step in the proposed method is classification of plant diseases based on leaf samples. The multi class SVM classifier is used to identify classes. Support Vector Machine (SVM) classifier is one of the dominant classification algorithm and gives better accuracy than other classification algorithms. The SVM creates the optimal hyper

plane dividing 2 classes. The hyper plane having the largest distance to the nearest data point of either class is considered as good separation. The nearest samples to the margin are chosen for deciding the hyper planes are called as support vectors. Basically the SVM was used for two-class problems. However, in real world it is more often to found the problems with multiple classes. In these situations, SVM can be extended from binary classifier to multi classifier with k classes. This extended form is known as Multi class SVM (m-SVM).

IV. CONCLUSION

This review paper discusses the contributions of multiclass Support Vector Machine (SVM) in the field of Agriculture. The most important characteristic of multi-class SVM is its ability to generalize well even with a small dataset. The selection of the feature extraction method and preprocessing of the data provide a higher contribution in the classification result. Early detection in plant disease is very important as it directly effects on crop productivity. A proper care can be taken if disease is detected as early as possible. Expertise are not always available and even they are costly. Hence the proposed method is very efficient,

REFERENCES

[1] S. S. Sannakki, V. S. Rajpurohit, V. B. Nargund, and P. Kulkarni, "Diagnosis and Classification of Grape Leaf Diseases using Neural Networks", IEEE 4th ICCCNT, 2013.

[2]. P. Chaudhary, A. K. Chaudhari, Dr. A. N. Cheeran and S.Godara, "Color Transform Based Approach for Disease Spot" International Journal of Computer Science and Telecommunications Volume 3, Issue 6, pp.65-70, June 2012

[3]. H. Al-Hiary, S. Bani-Ahmad, M. Reyalat, M. Braik, and Z. ALRahamneh, "Fast and Accurate Detection and Classification of Plant Diseases", IJCA, Vol-17, No.-1, pp. 31-38, March 2011.

[4]. Suman T. and Dhruvakumar T., "Classification of paddy leaf diseases using shape and color features", IJEEE, Volume 07, Issue 01, PP.239250, Jan- June 2015.

[5]. Ghaiwat Savita N, Arora Parul. Detection and classification of plant leaf diseases using image processing techniques: a review. Int J Recent Adv Eng Technol 2014;2(3):2347–812.

[6]. Arivazhagan S, Newlin Shebiah R, Ananthi S, Vishnu Varthini S. Detection of unhealthy region of plant leaves and classification of plant leaf diseases

[7]. Naikwadi Smita, Amoda Niket. Advances in image processing for detection of plant diseases. Int J Appl Innov Eng Manage 2013;2(11).

[8]. Patil Sanjay B et al. Leaf disease severity measurement using image processing. Int J Eng Technol 2011;3(5):297–301.

[9]. Chaudhary Piyush et al. Color transform based approach for disease spot detection on plant leaf. Int Comput Sci Telecommun 2012; 3(6).

[10].Bhanu B, Lee S, Ming J. Adaptive image segmentation using a genetic algorithm. IEEE Trans Syst Man Cybern Dec 1995;25:1543–67

[11].Vijayaraghavan Venkatesh, Garg Akhil, Wong Chee How, Tail Kang, Bhalerao Yogesh. Predicting the mechanical characteristics of hydrogen functionalized graphene sheets using artificial neural network approach. J Nanostruct Chem 2013;3:83.

[12].Kalaivanan, M., and K. Vengatesan. "Recommendation system based on statistical analysis of ranking from user." International Conference on Information Communication and Embedded Systems (ICICES), , pp. 479-484. IEEE, 2013

[13].P. Sanjeevikumar Vengatesan K, R. P. Singh, S. B. Mahajan," Statistical Analysis of Gene Expression Data Using Biclustering Coherent Column", International Journal of Pure and Applied Mathematics, Volume 114, Issue 9, Pages 447-454.

[14].P. Jaspreetkuar Sayyad Samee, Sarfaraz Khan, K. Vengatesan, Mahajan Sagar Bhaskar, P. Sanjeevikumar," Smart City Automatic Garbage Collecting System for a Better Tomorrow", International Journal of Pure and Applied Mathematics, volume 114, Issue 9,Pages: 455-463

[15] Pooja Kulinavar, Vidya I. Hadimani, " Classification of leaf disease based on Multiclass SVM classifier", International Journal of Advance Research, Ideas and Innovations in Technology, Volume3, Issue4,2017

[16] Shivaputra S.Panchal, Rutuja Sonar, "Pomegranate Leaf Disease Detection Using Support Vector Machine", International Journal Of Engineering And Computer Science ISSN: 2319-7242 Volume 5 Issues 6 June 2016, Page No. 16815-16818.