International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-6, April 2020

Automated Classification using SVM and Back Propogation Learning Technique

Shital Solanki, Ramesh Prajapati

***Abstract:** In this paper the comparative study of two supervised machine learning techniques for classification problems has been done. Due to the real-time processing ability of neural network, it is having numerous applications in many fields. SVM is also very popular supervised learning algorithm because of its good generalization power. This paper presents the thorough study of the presented classification algorithm and their comparative study of accuracy and speed which would help other researchers to develop novel algorithms for applications. The comparative study showed that the performance of SVM is better when dealing with multidimensions and continuous features. The selection and settings of the kernel function are essential for SVM optimality.*

***Keywords:** Artificial Neural Network, Generalization, Nonlinearity, Support vector machine, Supervised learning.*

I. INTRODUCTION

Classification problems are the problems which need to be identify the class or group of new data based on training of old data which are already classified under some class or group. SVM and ANN are well known supervised classification technique. In recent years, SVM is widely used in classification and recognition tasks. The Back-propagation neural network is best suited algorithm in case of multilayer feed forward NN. The Back-propagation neural network has very good application potential because of its easiness and accuracy. The main purpose of training a Neural Network is to produce the expected output for the given input set. [15]

II. SUPPORT VECTOR MACHINE

In recent years, SVM is one of the most powerful algorithms which are suited for both linear and nonlinear data. The original version of SVM presents the concept of maximal margin classifier introduced by Vapnik and Chervonenkis In 1963. In 1992, Vapnik has introduced kernalized version of SVM. This version is used to classify the data which is linearly nonseperable by adding kernel trick concept to SVM.

Revised Manuscript Received on April 30, 2020.
* Correspondence Author

Prof. Shital Solanki*, Assistant Professor, L.D.Engineering College, Ahmedabad,Gujarat Technological University,Gujarat ,India Email:sheetal172@gmail.com

Dr.Ramesh Prajapati, Assistant Professor, Indrashil Institute of Science & Technology, Rajpur- Kadi, Gujarat Technological University,Gujarat India. Email: tprajapati1984@gmail.com

Then after Vapnik has introduced soft margin classifier in SVM to allow misclassified data in 1995. [1]

A. CASE I: linearly separable Case

Consider a case of two classes of data which is strictly linearly separable.

Here SVM select two parallel hyper planes which separate the two classes of data. The distance between two hyper planes is known as margin.

The margin between two hyper planes as large as possible give better accuracy in classification Task.

The equation of separating hyper plane is given as:

$W.Yi + b = 0$ ⋮ (1)

Here, W is a weight vector

B is a bias

Yi is a input vector, i = 1, 2 . . . l

Define a vector y for the class, the classes are defined as±1

The equation of line which separates the classes is given as

$Yi.W + b \geq 1$ When J=+1 (2)

$Yi.W + b \leq -1$ When J=-1 (3)

The combined constraints can be written as:

$Ji(Yi.W + b) \geq 1\ \forall i$ i=1, 2 ...l (4)

The Margin from the separating line to the origin is:

M= $\frac{2}{w}$

The maximum margin is attained by minimizing norm of *w*.

$\frac{1}{2}\|W\|^2 st. Ji(Yi.W + b) \geq 1 \forall i$ (5)

In case, if there are infinite hyper plane who separate the data, with different values of W, need to find the optimal hyper plane which accurately classify as well as generalized for unseen data

B. CASE II: not fully linearly separable

In real world, not all data can be separated linearly, its having curved decision boundary. This kind of classification can be done by allowing some point as misclassified data using soft margin in SVM technique. It uses slack variables for misclassified data

The equation for the line dividing the classes will be given as:

$Yi.W + b + \sigma i \geq 1$ When J=+1 (6)

$Yi.W + b + \sigma i \leq -1$ When J=-1 (7)

The combined constraints can be written as:

$Ji(Yi.W + b) + \sigma i \geq 1\ \forall i$ i=1, 2, 3.....l (8)

$\frac{1}{2}\|W\|^2 + C\sum_{i=1}^{l} \sigma i \quad st. Ji(Yi.W + b) + \sigma i \geq 1\ \forall i$

i=1, 2, 3.....l (9)

Here C is used as a penalty parameter. The value of C measure under fitting and over fitting of training data. So taking the value of C too small leads to misclassification of training data. And too large value of C leads to poor generalization.

C. CASE III: Non-linear Data

In Real world, many datasets are nonlinear. But through nonlinear mapping dataset can be treated as linear dataset into a higher dimensional space. So the separating hyper plane considered as linear hyper plane in the higher dimensional feature space, but in actual it nonlinear input space. The kernel trick can be used to confine nonlinear boundaries. There are so many kernel functions exist which provide nonlinear mapping. The selection of best suited kernel depends on the nature of the data through trial and error.The linear kernel, polynomial kernel, RBF kernel and sigmoid kernel are more popular. [7].

III. BACKPROPOGATION NETWORK

Back-Propagation Neural Network (BPNN) is the most popular algorithm to train multilayer FFNN, proposed by Rumelhart[4] The BPNN algorithm backpropogate the error received at output to the whole NN to produce the expected output for the given input. Back-Propagation Neural Network is very flexible and having good learning capabilities. Due to these it has been successfully implemented in many applications [4]. BPNN uses Gradient descent approach, But the performance depends on so many parameter selected as initial Parameters. There exist different kinds of gradient descent approaches for BPNN algorithm to improve the training efficiency. Many researchers have used learning rate and momentum to improve the performance and convergence rate of NN. These parameters are often used in most of the approaches for weight adjustments [15]. A Back-Propagation network can have an input layer and an output layer. In between there is hidden layer. All units in each layer are connected to units in the other layers in feed forward fashion. A BPNN can be used to model any continuous function.[15] The input set is broadcasted to the network from input to the output layer. As a result of forward pass the actual output is generated. The expected outputs are used and given for training. The difference of actual and the expected outputs are used to calculate the error. Now the error is backpropogated through the network and the contribution of each neuron (processing units) for hidden layer and input layer are calculated.[15] Based on this calculation the corresponding adjustment needed for each processing unit in each layer have been calculated and as a result of backpropogation the connection weights are then adjusted. By iteratively updating weights in the network, performance is improved over time.[15]

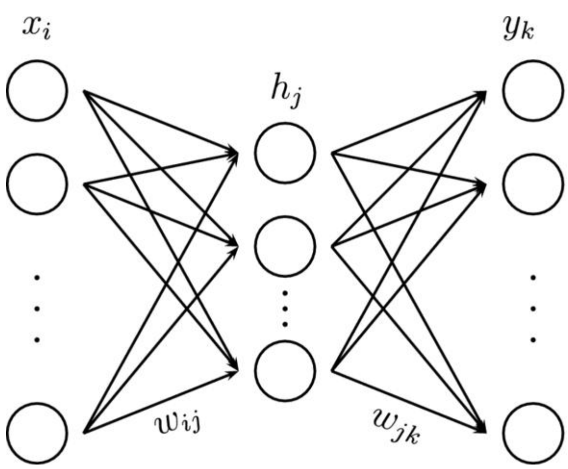

Figure 1: Feed forward Neural Network [19]

In this figure,

We assume i= 1, 2… K, j=1, 2…N and k= 1, 2…M,

Wij= Weights of connection from input to hidden layer

Wjk= Weights of connection from hidden to output layer

The output of an arbitrary neuron in the hidden layer h_j is given as

$$h_j = f(u_i) = f(\sum_{i=1}^{K} w_{ij}\, x_k) \qquad (1)$$

And similarly for the output of an arbitrary output neuron

$$y_k = f(u_k') = f(\sum_{j=1}^{N} w_{jk} h_j) \qquad (2)$$

The Objective function is root mean squared error summed over all elements in the output layer as

$$E = \frac{1}{2}\sum_{k=1}^{M}(y_k - t_k)^2 \qquad (3)$$

Where, tk's are target values for the output layer.

Rumelhart has introduced this back propagation method for training multi layer FFNN [4].

IV. LITERATURE REVIEW

A.Nonlinear SVM based on PSO [2]

Artificial neural network is widely popular approach to build rainfall forecasting model. Artificial neural network model works by minimizing the training error Rather than minimizing generalization error. In practical application ANN gives unpredictable output on noisy data.

SVM having good generalization power. Feature extraction again important step in rainfall forecasting model. Without efficient feature selection from input data model does not performed efficiently.

P.vikram developed rainfall forecasting model using support vector regression based on PSO.

PSO Algorithm which searches for best parameter for SVR [2]

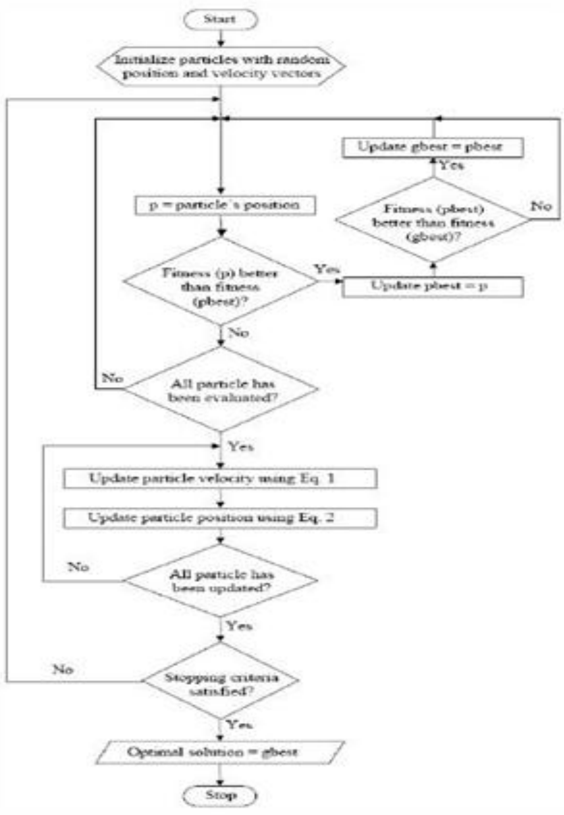

Figure2: PSO Flowchart [2]

Retrieval Number: C8350019320/2020©BEIESP
DOI: 10.35940/ijitee.C8350.049620
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Research Parameter selection is mostly done by trial and error procedure and then SVR model is build by considering different parameter set and best fitted model is selected among them by testing on validation set. If the parameter is not selected carefully it might be result in over fitting also. Here particle swam algorithm used to find optimal parameter for SVR. [1]

As per the experimental results the PSO based SVM model is better than BP–NN model for rainfall forecasting in terms of accuracy. [1]

B. Classification of Liver Cirhosis and Hemachromatosis [13]

Iron play a vital role in body functioning but excess amount of iron causes serious disease. Hemachromatosis is major liver disease, result of excess iron in liver. Stored iron can damage the organ and might result into Cirrhosis -chronic disease which represents the final phase of liver disease. Several Medical image classification schemes have been developed to help medical practitioners for the analysis of diseases. Some of them are ANN, SVM, K-Means and Nearest Neighbor. [13]

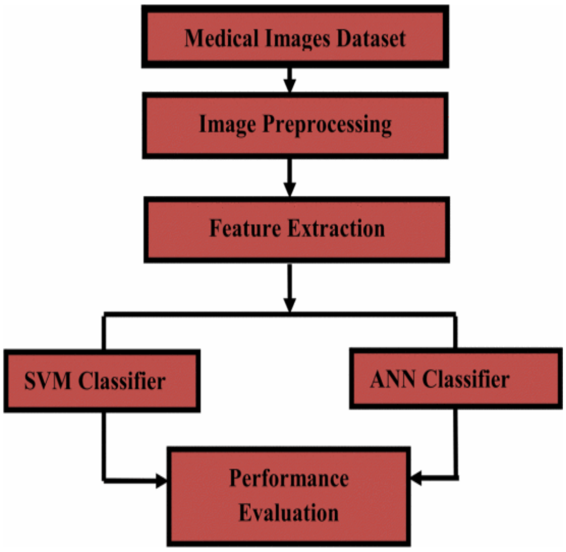

Figure 3: image classification scheme [13]

V. RESULTS AND DISCUSSION

Support vector Machine with linear kernel and ANN with 5-2-1 network structure used by author to emulate the accuracy of classification. Analysis of experimental results shown in table I and II.

Table- I: Confusion matrix of for SVM [13]

	Liver Cirrhosis	***Hemachromatosis***
Liver Cirrhosis	87 %	13%
Hemachromatosis	13%	87 %

Table- II: Confusion matrix of ANN [13]

	Liver Cirrhosis	***Hemachromatosis***
Liver Cirrhosis	67%	32%
Hemachromatosis	25%	75%

Based on the analysis result, SVM is having more classification accuracy than ANN.

VI. CONCLUSION

This paper presents a detailed study of SVM and BPNN algorithm. Various enhancements of both algorithms have been thoroughly reviewed. Based on Comparative study and analysis of experimental results, SVM is found better classification strength with accuracy of 87% than ANN in liver Disease.[13] PSO based SVM model is found better with greater forecasting accuracy[2] and can be used as an alternative model for rainfall forecasting.

Although BPNN algorithm is perform well in most of the Artificial Neural Network applications, convergence rate of network is slow and global minimum solution is not found. Generalization is not guaranteed even if the error is reduced to zero which include Over-fitting problem. For large dataset there are still chances of improvement in case of Back-propagation neural network.

Even though SVM outperform than the existing BPNN, there are still lot of scope of research which gives rise to further research in SVM. As a Future work, some efficient techniques may be used to select the initial parameters to enhance the performance of SVM.

REFERENCES

1. B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A Training Algorithm for Optimal Margin Classifiers,” in Proceedings of the Fifth Annual Workshop on Computational Learning Theory, 1992, pp. 144–152.
2. I. Technologies, “Rainfall forecasting using Nonlinear SVM based on PSO,” vol. 2, no. 5, pp. 2309–2313, 2011.
3. K. L. Hemalatha, H. Sneha, and H. R. Sumana, “Analysis and Design of Soft Clustering Method in Detection of Brain Tumor,” pp. 68–71, 2017.
4. Rumelhart, D. E., Hinton, G.E., Williams, R. J.: Learning Internal Representations by error Propagation. J. Parallel Distributed Processing: Explorations in the Microstructure of Cognition. (1986)
5. Lee, K., Booth, D., and Alam, P. A.: Comparison of Supervised and Unsupervised Neural Networks in Predicting Bankruptcy of Korean Firms. J. Expert Systems with Applications. 29, 1- -16 (2005)
6. Chun Lu, Bingxue Shi and Lu Chen, “An expandable on-chip Back Propagationlearning neural network chip”, Int. J.Electronics, Vol 90, No. 5, 331-340, 2003.
7. Zweiri, Y.H., Seneviratne, L D., Althoefer, K.: Stability Analysis of a Three-term Back-propagation Algorithm. J. Neural Networks. 18, 1341- -1347 (2005)
8. Jian-Pei Zhang, Zhong-Wei Li, and Jing Yang, “A parallel SVM training algorithm on large-scale classification problems,” 2005 Int. Conf. Mach. Learn. Cybern., no. August, p. 1637–1641 Vol. 3, 2005.
9. I. M. Gorovyi and D. S. Sharapov, “Comparative analysis of convolutional neural networks and support vector machines for automatic target recognition,” MRRS 2017 - 2017 IEEE Microwaves, Radar Remote Sens. Symp. Proc., pp. 63–66, 2017.
10. V. J. Prakash and D. L. M. Nithya, “A Survey On Semi-Supervised Learning Techniques,” Int. J. Comput. Trends Technol., vol. 8, no. 1, pp. 25–29, 2014.
11. B. Sanjaa and E. Chuluun, “Malware detection using linear SVM,” 8th Int. Forum Strateg. Technol. 2013, IFOST 2013 - Proc., vol. 2, pp. 136–138, 2013.
12. S. Alam, M. Kang, J. Y. Pyun, and G. R. Kwon, “Performance of classification based on PCA, linear SVM, and Multi-kernel SVM,” Int. Conf. Ubiquitous Futur. Networks, ICUFN, vol. 2016–Augus, pp. 987–989, 2016.
13. O. D. Fenwa, F. A. Ajala, and A. M. Aku, “Performance evaluation of support vector machine and artificial neural network in the classification of liver cirhosis and hemachromatosis,” Proc. - Int. Conf. Comput. Vis. Image Anal. Appl. ICCVIA 2015, pp. 1–6, 2015.

14. D. Nayak, A. Mahapatra, and P. Mishra, "A Survey on Rainfall Prediction using Artificial Neural Network," Int. J. Comput. …, vol. 72, no. 16, pp. 32–40, 2013.
15. Shital Solanki,H.B.Jethva Modified Back Propagation Algorithm of Feed Forward NetworkInternational Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-2, Issue-6, May 2013
16. "Advances in Natural Computation", Springer Science and Business Media LLC,2005

AUTHORS PROFILE

Shital Solanki is presently working as an Assistant Professor, at L.D. Engineering College, Ahmadabad in Information Technology Department. She has done her Master in Information Technology in year 2013.She is pursuing her Ph.D. in Computer/IT Engineering from Gujarat Technological University Ahmadabad, Gujarat. She is having 18 years of teaching experience. Her area of interest is in Machine Learning & Artificial Intelligence. She is Lifetime Member of Indian Society of Technical Education.

Dr. Ramesh Prajapati is presently working as an HoD & Assistant Professor, Indrashil University, Rajpur, Kadi, india in CE Department and He is having more than 12 years of teaching experience. He has done his Ph.D. in Computer/IT Engineering from Rai University,Ahmadabad Gujarat. His area of interest is in cloud computing, Machine Learning,Artificial Intelligent and networking.