International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 5 Issue: 9 115 – 120

A Survey on Hybrid Techniques Using SVM

KolluruVenkata Nagendra
Research Scholar,
Department of Computer Science
Vikrama Simhapuri University.
Nellore, Andhra Pradesh, India
kvnscholar@gmail.com

Dr. Maligela Ussenaiah
Assistant Professor,
Department of Computer Science,
Vikrama Simhapuri University.
Nellore, Andhra Pradesh, India

Abstract: Support Vector Machines (SVM) with linear or nonlinear kernels has become one of the most promising learning algorithms for classification as well as for regression. All the multilayer perceptron (MLP),Radial Basic Function(RBF) and Learning Polynomials are also worked efficiently with SVM. SVM is basically derived from statistical Learning Theory and it is very powerful statistical tool. The basic principal for the SVM is structural risk minimization and closely related to regularization theory. SVM is a group of supervised learning techniques or methods, which is used to do for classification or regression. In this paper discussed the importance of Support Vector Machines in various areas. This paper discussing the efficiency of SVM with the combination of other classification techniques.

KeyWords: *Rapid Eye time series, Space-mapped Binary Tree SVM, Linear Kernel SVM .*

*****

1. INTRODUCTION

Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data set. These tools can include statistical models, mathematical algorithm and machine learning methods. Consequently, data mining consists of more than collection and managing data, it also includes analysis and prediction. Classification technique is capable of processing a wider variety of data than regression and is growing in popularity. There are several applications for Machine Learning (ML), the most significant of which is data mining. People are often prone to making mistakes during analyses or, possibly, when trying to establish relationships between multiple features. This makes it difficult for them to find solutions to certain problems. Machine learning can often be successfully applied to these problems, improving the efficiency of systems and the designs of machines. Numerous ML applications involve tasks that can be set up as supervised. SVM have been used to classify data in various domains like land cover classification, species distribution, medical binary classification, fault diagnosis, character classification, speech recognition, radar signal processing, habitat prediction etc... In the present paper, we have concentrated on the SVM technique. In particular, this work is concerned with classification problems using SVM and the combination of SVM with other classification techniques.

Support Vector Machines is presented in below section. Section 3 described SVM with remote sensing area. In pattern recognition and prediction the efficiency of SVM discussed in section 4 and 5. Detailed analysis of SVM with other methods described in section6. Table –I shows that the SVM work in different classification techniques in various fields. Finally, the last section concludes this work.

II. OVERVIEW OF SVM

SVM is used to construct a hyper-plane from a set of data vectors whose class labels are known in advance which can be used to classify data vectors and also to assign class labels to unlabelled data. Data sets are mapped onto a dot product space via a linear kernel function that is of the form: K(xi,xj)={f(xi),f(xj)}. Later, the decision function is stated in terms of kernel function as mentioned in equation (1):

$$f(x) = \sum_{j=1}^{Sv} \alpha j y j K(x, xj) + c \qquad (1)$$

C is a user-defined aspect which maintains the balance between the number of non-separable points and the machine complexity and Sv denotes the number of support vectors used for which 0≤ ≤C while is Lagrange multipliers. Bias c is a scalar which is computed by using any support vector.Then, the optimal hyper-plane corresponds to f(x)=0. Test data are classified as mentioned in equation (2):

$$x \in \begin{cases} \text{positive class if } f(x) > 0 \\ \text{negative class if } f(x) < 0 \end{cases} \quad (2)$$

Once the separating hyper-plane is defined we can use the same for assigning class labels to unlabelled data sets. Initially the feature vectors and the labels are observed for each sample item in the data set. The classifier aims at finding a function f :x→{±1} which assigns a class label y to an unlabelled sample data item x[7].

Thus, the separation learned from the limited training data set can be applied later to validate against the testing data and later to assign class labels to an unlabelled data set. The training set comprising of instance label pairs is pointed out in equation (3).

$$(x, y) = \{ (x_1,y_1), (x_2,y_2), \ldots\ldots\ldots (x_n,y_n)\} \quad (3)$$

Where and {-1, +1} and SVM needs solution of the optimization issue which is to maximize the margin between contrasting data items and is denoted by

$$\frac{1}{2}W^TW + \sum_{i=1}^{l} \xi_i C \text{ Subject to as given in equation (4)}$$

$$y_i (w^T\Phi(x_i) + b) \geq 1-\xi_i\xi_i \geq 0 \quad (4)$$

Here training vector is mapped to infinite dimensional space, then SVM finds the linear separating hyper-plane with the maximal margin in the higher dimensional space where C>0

III. SVM IN THE AREA OF REMOTE SENSING

In agricultural resource monitoring, the crop mapping is one major component and it uses remote sensing. At object level in irrigated landscapes, crop classification is done by SVM. The total 71 multi seasonal spectral and geo statistical features computed from RapidEye time series as a Input to the Classification. SVM is a well-established machine learning technique that has not only given promising accuracies in crop classification, but also was shown to give soft outputs that can be interpreted as a measure of uncertainty in classification. Unlike the other algorithms the soft output from SVM in the context of land use mapping has not attained much attention yet. On the small training data sets the SVM can perform efficient classification. In remote sensing applications SVM were shown to perform more accurately than other classifiers in land cover and crop classification.

Fish images were acquired and sent by smart phone, and the method utilized was comprised of the Color and texture sub images of fish skin were obtained from original images. Color features, statistical texture features and wavelet-based texture features of the color and texture sub images were extracted, and six groups of feature vectors were composed.

Classifications of different water levels facilitate in preserving water reserves and maintain the equilibrium in the ecosystem. A powerful statistical tool called support vector machines is used to classify the said drainage water remote sensed spatial data sets. To boost the performance of support vector machines classifier a new generic algorithm based on parametric distribution model will be proposed. Later several evaluation metrics like kappa statistics are used to compare the results of the proposed algorithm with multi-layer perceptron neural networks and naive bayes classifiers.

IV. SVM IN THE AREA OF PATTERN RECOGITION

On the basis of analysis of the pattern recognition theory and traditional classification method of birds, we mainly study on the extraction and selection of images feature, through extracting the birds' pictures feature and making some mathematical transformation analysis, we select the most representative classification feature of birds' species. By introducing the image analysis and processing of birds into the research of birds' recognition and classification, this aims to explore the establishment of birds' automatic classification mechanism which is based on color images. A careful study of support vector machines (SVM) principle and kinds of multiply classification methods are used. A kind of multiply classification system of SVM decision tree with priori information is used for pattern recognition. Finally, through selecting the birds characteristics trained, and selecting the test set to do experiment by this SVM classification, a better recognized result can be gotten.

V. SVM IN THE AREA OF PREDICTION

Classification is a data mining technique that is used to construct a model and allocate class labels to data instances. It is also used to predict class labels for unlabelled data. In recent years, Classification and prediction of rainfall has gained a lot of significance due to frequent recurrence of hurricanes and floods in the recent past. There is a large amount of data that has been accumulated pertaining to rainfall over a period of time. This huge voluminous data is tossing challenges to the research community in terms of processing and active analysis. Current methods and algorithms are insufficient to do effective analysis. Hence advanced data mining techniques like machine learning

algorithms and their hybridization techniques are suitable for processing and effective analysis.

The Researchers discussed different classification technique; Linear Kernel based Support Vector Machines for classifying spatial data. Performance Analysis of these algorithms based on various parameters is carried out and it is found that performance of SVM based classification is very good for the variety of datasets used. A novel approach to classify remote sensed data using a Hybrid SVM classifier is very efficient. Support Vector machines and Neural Network methodologies are used to construct the hybrid classifier for effective spatial data analysis. The proposed hybrid method when applied to classify the rainfall data in a any region and it is found that it gives better results when compared to traditional neural network and support vector machine classification methods used individually without any compromise in classification accuracy.

VI. LITERATURE SURVEY AND ANALYSIS OF SUPPORT VECTOR MACHINES

There have been many researchers who have attempted SVM for classification. These classification metods were combined with SVM and they gained more accurate results. The Hybrid SVM will gives the efficient results while compared with the single SVM. The works in [1],[2],[3],[4],[5],[6] have designed SVM for remote sensed data. The works in [7],[8],[9],[10],[11],[12],[13],[14], have focused on Pattern recognition techniques, whereas the works in [18],[19],[20],[21] have focused on Binary Tree SVM.

F.Low and U.Michel [22] study Support Vector Machines (SVM) are used for crop classification in irrigated landscapes at the object-level. Input to the classifications is 71 multi seasonal spectral and geostatistical features computed from RapidEye time series. The random forest (RF) feature importance score was used to select a subset of features that achieved optimal accuracies. The relationship between the hard result accuracy and the soft output from the SVM is investigated by employing two measures of uncertainty, the maximum a posteriori probability and the alpha quadratic entorpy.

In [23], the researchers remote diagnose of fish diseases for farmers is unrealized in China, but use of mobile phones and remote analysis based on image processing can be feasible due to the widespread use of mobile phones with camera features in rural areas. This paper presents a novel method of classifying species of fish based on color and texture features and using a multi-class support vector machine (MSVM). Fish images were acquired and sent by smart phone, and the method utilized was comprised of the following stages. Color and texture sub images of fish skin were obtained from original images. Color features, statistical texture features and wavelet-based texture features of the color and texture sub images were extracted, and six groups of feature vectors were composed.

In [24], authors applied SVM which is a very powerful statistical tool compared with K-nearest Neighbor classifier based on some parameters. The research establishes that the SVM approach of classification is more efficient as compared to KNN because SVM provides higher values. In [25], authors applied Euclidean SVM approach for oil and gas pipeline failure prediction system using long range uiltrasonic transducers.

Baoping Tang and Tao Song[26], discussed a novel fault diagnosis method, manifold learning combined with the shanon wavelet vector machine, which was proposed for wind turbine transmission systems. It has taken the advantage of manifold learning algorithm with good performance in non linear dimensionaity reduction and SWSVM.

Table I presents categorization of different approaches of SVM with other classification methods. The categorization is based on the different fields and various areas which it can be applied to find the efficiency of SVM.

TABLE I. ANALYSIS OF DIFFERENT APPROACHES OF SVM WITH OTHER METHODS

Researcher	Region (Global/Local)	Technique Used with SVM	Field	Applied on
F.Low,U.Michel and S.Dech	Global	RapidEye Timeseries	Remote sensing	Crop Classification
Jing Hu,Daoliang Li and YueqiHan	Global	Multi Class SVM	Image Processing	Fish Species by color, Texture

 ISSN: 2321-8169
115 – 120

Xiaowei Yang and Qiaozhen Yu	Global	Binary Tree SVM	Pattern Recognition and Machine Learning	Multi Class Classification
Yang Shao and Ross Slunetter	Global	MODIS Time Series Data (MODerate resolution Imaging Spectro radiometer)	Image Processing	Land Cover Characterization
Tarun Rao and N.Rajasekhar	Global (India)	Linear Kernel	Image Processing (Remote sensed Data)	Fish Species
Hong Ji Li,Han Lin and Weibin Chen	Global (China)	SVM Decision Tree	Digital Graphic Recognition	Bird Species
AD Deleep and C.Chandrasekhar	Global (India)	Dynamic Kernel based SVM	Speech Recognition	Length patterns of long duration speach
T.Rajanikanth and T.Rao	Global (India)	Kappa Statistics	Drainage water level Spatial data	Drainage water level Classification
Lam Hon Lee,Raj Prasad and Dino Isa	Global	Ultrasonic transducers and Euclidean SVM	Non Destructive Testing(NDTs) method based on Long Range Ultrasonic Transducers(LRUTs)	Oil and Gas pipeline failure prediction System
Xinjun Peng and Dong Xu	Global	TPMSVM Twin Parametric Margin SVM	Structural Granularity	Binary Classification
MiaoLiu and DuoLi	Global (China)	Random forecast (RF) and Classification And Regression Tree (CART)	E-Tongue	Ensemble Learning
Till Rumpf and Martin Weis	Global	Sequential SVM	Image processing and Classification	Small grain weed species
Hassiba Nemmous and Youcef Chibani	Global	Multiple SVM	Image processing	Land cover change Detection
Daniel J. Sebald and James A. Bucklew	Global	Inter Symbol Interference (ISI)	Nonlinear Equalization	addressing nonlinearities in communication channels exhibiting ISI
	Global		The smoothing	large-scale complementarily

Tie Ni a and Jun Zhai b	(China)	Karush-Kuhn-Tucker (KKT) systems	Newton algorithm	problems and optimization problems
Tiemen Strobbea, Francis wyffelsb and Ruben Verstraetena	Global (China)	one-class support vector machines (SVMs) with graph kernels	architectural designs belonging	a particular architectural style or corpus
M. Malvoni, M.G. De Giorgi and P.M. Congedo	Local (India)	the principal component analysis (PCA) ,Least Squares Support Vector Machines (LS-SVM)	Photovoltaic forecast based on hybrid PCA-LS SVM using dimensionality reducted data	Photovoltaic forecast
Engin Esme and Bekir Karlik	Local (India)	Unsupervised fuzzy clustering c-mean (FCM)	To recognize patterns of perfume	Clustering

The objective of the present work is to find the efficient usage of SVM in different fields and also explained the benefits of combination of hybrid SVMs. With this survey we finalized that SVM is good for classification.

VIII. CONCLUSION

In this Paper the survey yields that the Performance of SVM based classification is very good for the variety of datasets used. A hybrid approaches to classify data using a SVM classifier is very efficient. Support Vector machines was combined with all the other classification techniques in different fields like pattern recognition, Image processing, digital graphical recognition, Euclidean SVM leads the powerful results. In this paper the SVM produce effective and efficient results with the combination of remaining classification methods.

REFERENCES

[1] Balaguer, A., Ruiz, L., Hermosilla, T., Recio, J., 2010. Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. Computers & Geosciences 36 (2), 231–240.

[2] Bazi, Y., Melgani, F., 2006. Toward an optimal SVM classification system for hyperspectral remote sensing images. IEEE Transactions on Geoscience and Remote Sensing 44 (11), 3374–3385.

[3] Bloch, I., 1996. Information combination operators for data fusion: A comparative review with classification. IEEE Transactions on Systems, Man, and Cybernetics- Part A: Systems and Humans 26 (1), 52–67.

[4] Broge, N., Leblanc, E., 2000. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. Remote Sensing of Environment 76 (2), 156–172.

[5] Chen, J., 1996. Evaluation of vegetation indices and a modified simple ratio for boreal applications. Canadian Journal of Remote Sensing 22 (3), 229–242.

[6] Justice, C.O., Becker-Reshef, I., 2007. Report from the Workshop on Developing a Strategy for Global Agricultural Monitoring in the Framework of Group on Earth Observations (GEO). Rome.

[7] Arribas, J.I., Sanchez-Ferrero, V., Ruiz-Ruiz, G., Gomez-Gil, J., Gomez-Gil, J., 2011. Leaf classification in sunflower crops by computer vision and neural networks. Computers and Electronic in Agriculture 78 (1), 9–18.

[8] Bermejo, S., 2007. Fish age classification based on length, weight, sex and otolith morphological features. Fisheries Research 84 (2), 270–274.

[9] Bottou, L., Cortes, C., Denker, J., Drucker, H., Guyon, I., Jackel, L., LeCun, Y., Muller, U., Sackinger, E., Simard, P., Vapnik, V., 1994. Comparison of classifier methods: a case study in handwriting digit recognition. In: Proceedings of International Conference on Pattern Recognition, Jerusalem, pp. 77–87.

[10] Camargo, A., Smith, J.S., 2009. Image pattern classification for the identification of disease causing agents in plants. Computers and Electronics in Agriculture 66 (2), 121–125.

[11] Chang, C.C., Lin, C.J., 2001. LIBSVM: a library for support vector machines. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

[12] Chen, J., Wang, C., Wang, R.S., 2009. Adaptive binary tree for fast SVM multiclass classification. Neuro computing 72 (13–15), 3370–3375.

[13] Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J.L., 2001. Color image segmentation: advances and prospects. Pattern Recognition 34 (12), 2259–2281.

[14] Coifman, R.R., Wickerhauser, M.V., 1992. Entropy-based algorithms for best basis selection. IEEE Transactions on Information Theory 38 (2), 713–718.

[15] Bogdanov, A.V., Schöner, G., Steinhage, A., Sandven, S., 2003.Multiple classifier system based on attractor dynamics. Proc. International Geoscience and Remote Sensing Symposium,IGARSS/IEEE, Toulouse, 24–28 July, pp. 3635–3637.

[16] Brown, M., Lewis, H.G., Gunn, S.R., 2000. Linear spectral mixture models and support vector machines for remote sensing. IEEE Transactions on Geoscience and Remote Sensing 38 (5),2346–2360.

[17] Bruzzone, L., Melgani, F., 2003. Classification of hyper-spectral images with support vector machines: multi-class strategies. Proc. Image and Signal Processing for Remote Sensing IX, SPIE, Barcelona, 8–9 September, pp. 408–419.

[18] Belousov, A.I., Verzakov, S.A., Von Frese, J., 2002. A flexible classification approach with optimal generalisation performance: support vector machines. Chemometrics and Intelligent Laboratory Systems 64, 12–55.

[19] X. Peng, D. Xu, Robust minimum class variance twin support vector machine classifier, Neural Computing and Applications (2013), http://dx.doi.org/10.1007/s00521-011-0791-3.

[20] X. Peng, D. Xu, Twin Mahalanobis distance-based support vector machines for pattern recognition, Information Sciences 200 (1) (2013) 22–37.

[21] Kima, J.S., Hong, K.S., 2009. Color–texture segmentation using unsupervised graph cuts. Pattern Recognition 42 (5), 735–750.

[22] F.Low , U.Michel and S.Dech, "Impact of feture selection on the accuracy and spatial uncertainty of per –field crop classification using Support Vector Machines", ISPRS journal of photogrammetric and remote sensing -2013.

[23] Jing Hu,Daoling Li and Yueqi Han, " Fish Species classification by color, texture and multi class support vector machine using computer vision", Computer and electrical in agriculture-2012

[24] Tarun Rao,N.Rajaekhar,T.V.Rajanikanth and Sundar KS, "Classification of Remote Sensed Data using Liner Kernel Based Support Vector Machines", ICCC Dec2013 Trivendrum.

[25] Lam Hong Lee,Rajprasad, Lai Hung Lo and Dino Isa," Oil and Gas Pipeline failure prediction System using long ultrasonic transducers and Euclidean Support Vector Machines classification approach",Expert system applications-2013.

[26] Baoping Tang,Taosong,Feng Li and Lei Deng," Fault diagnosis a wing turbine transmission system based on manifold leaning and Shannon wavelet support vector machine", Renewable energy -2013 Elsevier.

[27] Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 121–167.Carlotto, M.J., 1997. Detection and analysis of change in remotely sensed imagery with application to wide area surveillance. IEEE Transactions on Image Processing 6 (1), 189–202.

[28] Chibani, Y., Nemmour, H., 2003. Kalman filtering as a multilayer perceptron training algorithm for detecting changes in remotely sensed imagery. Proc. International Geoscience and Remote Sensing Symposium, IGARSS/IEEE, Toulouse, 21–25 July, pp. 4101–4103.

[29] Cho, S.B., 1995. Fuzzy aggregation of modular neural networks with ordered weighted averaging operators. International Journal of Approximate Reasoning 13 (4), 359–375.

[30] Cho, S.B., 2002. Fusion of neural networks with fuzzy logic and genetic algorithm. Integrated Computer-Aided Engineering, vol. 9. IOS Press, pp. 363–372.

[31] Cho, S.B., Kim, J.H., 1995. Combining multiple neural networks by fuzzy integrals for robust classification. IEEE Transactions on Systems, Man, and Cybernetics 25 (2), 380–384.

BIOGRAPHIES

K. Venkata Nagendra, working as Assistant Professor in the Department of Computer Science Engineering at Geetanjali Institute of Science and Technology, Nellore, Andhrapradesh, India. He has 6 years of experience in the field of teaching. He is a research scholar in Vikrama Simhapuri University. He did his M.Tech in ANU, Guntur. His areas of interests are Data warehousing and Data Mining and Cloud Computing.

Dr. Maligela Ussnaiah, working as Assistant Professor in the Department of Computer Science in Vikrama Simhapuri University, Nellore, Andhra Pradesh, India. He is having 8 years of teaching experience. He did his PhD in computer science from SriKrishna Devaraya Univeristy, Ananthapur, Andhra Pradesh. His areas of interests are Networks, Mobile Wireless Networks, Data warehousing and Data Mining and Image processing.