International Journal of Database Theory and Application
Vol.8, No.1 (2015), pp.169-186
http://dx.doi.org/10.14257/ijdta.2015.8.1.18

A Comprehensive Survey on Support Vector Machine in Data Mining Tasks: Applications & Challenges

Janmenjoy Nayak, Bighnaraj Naik* and H. S. Behera

Department of Computer Science Engg. & Information Technology
Veer Surendra Sai University of Technology,
Burla - 768018, Odisha, India
*mailforjnayak@gmail.com, *mailtobnaik@gmail.com,*
mailtohsbehera@gmail.com

Abstract

During the last two decades, a substantial amount of research efforts has been intended for support vector machine at the application of various data mining tasks. Data Mining is a pioneering and attractive research area due to its huge application areas and task primitives. Support Vector Machine (SVM) is playing a decisive role as it provides techniques those are especially well suited to obtain results in an efficient way and with a good level of quality. In this paper, we survey the role of SVM in various data mining tasks like classification, clustering, prediction, forecasting and others applications. In broader point of view, we have reviewed the number of research publications that have been contributed in various internationally reputed journals for the data mining applications and also suggested a possible no. of issues of SVM. The main aim of this paper is to extrapolate the various areas of SVM with a basis of understanding the technique and a comprehensive survey, while offering researchers a modernized picture of the depth and breadth in both the theory and applications.

Keywords: *Support Vector Machine (SVM), Data Mining, Artificial Neural Network (ANN)*

1. Introduction

In recent years, the Artificial Neural Networks (ANNs) have been playing a significant role for variants of data mining tasks which is extensively popular and active research area among the researchers. The intend of neural network is to mimic the human ability to acclimatize to varying circumstances and the current environment. Starting from Mc Culloch-Pitts network, the research is highly popular to some of the higher order neural network. The methodology of an Artificial Neural Network , intended to imitate few capabilities of the human brain and has demonstrated great prospective for various low level computations and embodies prominent features such as learning, fault tolerance, parallelism *etc*. ANN is a popular technique for the application in most of the data mining fields including classification [1-3], forecasting [4-12], functional approximation [13-15], rule extraction [16-19], pattern recognition and medical applications [20-22]. In the present day of research, ANN has stood forward as a strong alternative to the traditional recognition models. The computer science research is already in young age for the implementation technique of some popular ANN models like Hopfield network, Multilayer Perceptron, Self Organizing Feature Map, Learning Vector Quantization, Radial Basis Function, Cellular Neural Network, Adaptive Resonance Theory Networks, Counter Propagation Networks, Back Propagation Network and Support Vector Machines *etc*. Due to the presence of all these neural networks, it is reasonable to expect a rapid increase in our understanding of artificial neural networks leading to improved network paradigms and a host of application opportunities.

The insidious use of Support Vector Machine (SVM) in various data mining applications makes it an obligatory tool in the development of products that have implications for the human society. SVMs, being computationally powerful tools for supervised learning, are widely used in classification, clustering and regression problems. SVMs have been successfully applied to a variety of real-world problems [23] like particle identification, face recognition, text categorization, bioinformatics, civil engineering and electrical engineering *etc*.

In this work, a detailed survey is performed in almost majority of the data mining fields starting from the year 2001 to 2014. Section 2 gives the fundamental ideas about the Support Vector Machines, Section 3 outlines a road map to SVM, Section 4 gives the statistical analysis about SVM research and Section 5 concludes our work with some of the major issues of SVM technique. The majority of the work cited in this paper is journal articles. The reason for this is that, we want to report on data mining applications that are extensibly used through SVM. Specifically, we have considered papers written in the English language and are of Journals and Conference proceedings types. This paper has two major objectives which are to provide a short background to the SVMs that are relevant to data mining tasks and Provide a review of the state of the art in the application of SVM methods in data mining.

2. Support Vector Machine

Support Vector Machines (SVMs) as originally proposed by Vladimir Vapnik[24] within the area of statistical learning theory and structural risk minimization, have demonstrated to work successfully on various classification and forecasting problems. SVMs have been used in many pattern recognition and regression estimation problems and have been applied to the problems of dependency estimation, forecasting and constructing intelligent machines [25]. SVMs have the prospective to capture very large feature spaces, due to the generalization principle which is based on the Structural Risk Minimization Theory (SRM) *i.e.*, the algorithm is based on guaranteed risk bounds of statistical learning theory[26].

In MLP classifiers, the weights are updated during the training phase for which the total sum of errors among the network outputs and the desired output is minimized. The performance of the network strongly degrades for small data sizes, as the decision boundaries between classes acquired by training are indirect to resolute and the generalization ability is dependent on the training approach. In contrast to this, in SVM the decision boundaries are directly determined from the training data set for which the separating margins of the boundaries can be maximized in feature space.

A SVM is a maximum fringe hyperplane that lies in some space and classifies the data separated by non-linear boundaries which can be constructed by locating a set of hyperplanes that separate two or more classes of data points. After construction of the hyperplanes, the SVM discovers the boundaries between the input classes and the input elements defining the boundaries (support vectors [27]). From a set of given training samples labeled either positive or negative, a maximum margin hyperplane splits the positive or negative training sample, as a result the distance between the margin and the hyperplane is maximized. If there exist no hyperplanes that can split the positive or negative samples, a SVM selects a hyperplane that splits the sample as austerely as possible, while still maximizing the distance to the nearest austerely split examples.

Figure 1 indicates a linearly separable hyper plane, where there are two groups of data points represented by '$\otimes$' and 'Δ'. There may be possibility of an infinite no. of hyper planes but in the described figure, only one hyper plane represented by solid line optimally separates the sample points and is situated in between the maximal margins.

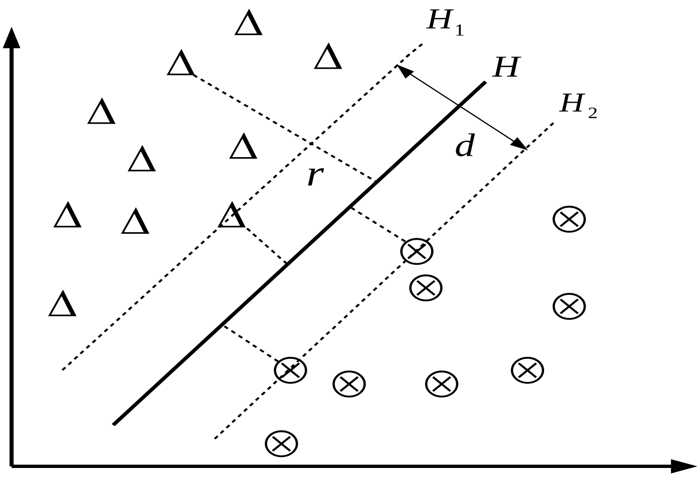

Figure 1. Linearly Separable Samples Indicated in a Hyperplane

Suppose we have N training samples like $\{(p_1, q_1), (p_2, q_2)...(p_N, q_N)\}$ where $p_i \in R^d$ and $q_i \in \{1, -1\}$. Equation (1) represents the equations of a hyper plane used for data portioning in SVM.

$$W \cdot p + b = 0 \tag{1}$$

where W is a weight, p is the training sample and b is the bias of the hyper plane. The margin between two classes are to be maximized, for which $|W|$ should be minimized [28] subject to the condition

$$q_i(W * p_i + b) \geq 1 \tag{2}$$

The optimization problem can be defined as

$$\min_{W,b} \frac{1}{2} W^T W \text{ with respect to } q_i(W * p_i + b) \geq 1 \text{, for } i = 1,2,... N \tag{3}$$

for the minimization of the value $\frac{1}{2} W^T W$, p_i and $q_i \in \{1, -1\}$.

By introducing the Lagrange multiplier $\beta_1, \beta_2, ... \beta_N \geq 0$ for solving this problem,

$$L(W, b, \beta) = \frac{1}{2} W^T W - \sum_{i=1}^{N} \beta_i q_i (W * p_i + b) + \sum_{i=1}^{N} \beta_i \tag{4}$$

Hence, the problem becomes [29] $\max\ L(\beta) = \sum_{i=1}^{N} \beta_i - \frac{1}{2} \sum_{i,j=1}^{N} \beta_i \beta_j q_i q_j p_i \cdot p_j$ subject to

$$\sum_{i=1}^{N} \beta_i q_i = 0, \quad \beta_i \geq 0 \tag{5}$$

Now, if β_i^* is an optimal solution, the corresponding optimal bias and weight values can be updated as;

$$W^* = \sum_{i=1}^{N} q_i \beta_i^* p_i \tag{6}$$

$$b^* = -\frac{1}{2} W^* [p_m + p_n]$$

where p_m, p_n are support vectors.

For linearly non-separable data (Figure 2) we introduce a non-negative variable called slack variable $V_i \geq 0$

Now, equation (2) becomes $q_i(W * p_i + b) + V_i \geq 1$ (7)

So, equation (3) will have the form,

$$\min_{W,b,V_i} \frac{1}{2} W^T W + R\left(\sum_{i=1}^{N} V_i\right) \tag{8}$$

subject to $q_i(W * p_i + b) + V_i - 1 \geq 0$

where, $V_i \geq 0$, i = 1, 2, ..., N and R is a positive parameter [30] if the sample p_i is in the correct region $i.e.$ $0 \leq V_i < 1$ otherwise, $V_i >$ 1.

In the same way by introducing Lagrange multiplier $\beta_1, \beta_2, ..., \beta_N \geq 0$. For solving the problem

$$\max_{\beta} \; L(\beta) = \sum_{i=1}^{N} \beta_i - \frac{1}{2} \sum_{i,j=1}^{N} \beta_i \beta_j q_i q_j p_i \cdot p_j$$

Subject to $\sum_{i=1}^{N} \beta_i * q_i = 0$, $0 \leq \beta_i \leq R$, where N represents the number of support vectors.

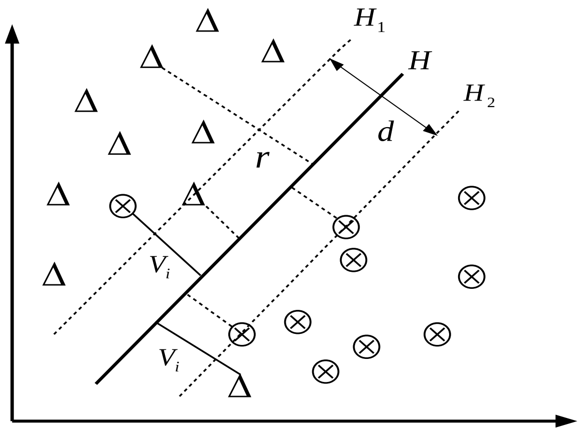

Figure 2. Linearly Non-Separable Samples Indicated in a Hyperplane

3. Literature Review

In this section, let us survey some major contributions towards SVM and its successful applications in various data mining tasks.

R. Burbidge *et. al.*, [31] have shown that the support vector machine (SVM) classification algorithm, proves its potential for structure–activity relationship analysis. In a benchmark test, they compared SVM with various machine learning techniques currently used in this field. The classification task involves in predicting the inhibition of dihydrofolate reductase by pyrimidines, using the data obtained from the UCI machine learning repository. Among three tested artificial neural networks, they found that SVM is significantly better than all of these.

Giorgio Valentini [32] have proposed classification methods, based on non-linear SVM with polynomial and Gaussian kernels, and output coding (OC), ensembles of learning machines to separate normal from malignant tissues, to classify different types of lymphoma and to analyze the role of sets of coordinately expressed genes in carcinogenic processes of lymphoid tissues. By using gene expression data from ''Lymphochip'', he has shown that SVM can correctly separate the tumoural tissues, and OC ensembles can be successfully used to classify different types of lymphoma.

Shutao Li *et. al.*, [33] have applied SVMs by taking DWFT as input for classifying texture, using translation-invariant texture features. They used a fusion scheme based on simple voting among multiple SVMs, each with a different setting of the kernel parameter, to alleviate the problem of selecting a proper value for the kernel parameter in SVM training and performed the experiments on a subset of natural textures from the Brodatz album. They claim that, as compared to the traditional Bayes classier and LVQ, SVMs, in general, produced more accurate classification results.

A training method to increase the efficiency of SVM has been presented by Yiqiang Zhan [34] for fast classification without system degradation. Experimental results on real prostate ultrasound images show good performance of their training method in discriminating the prostate tissues from other tissues and they claim that their proposed training method is able to generate more efficient SVMs with better classification abilities.

Yuchun Tang *et. al.*, [35] have developed an innovative learning model called granular support vector machines for data classification problems by building just two information granules in the top-down way. The experiment results on three medical binary classification problems show that granular support vector machines proposed in their work provides an interesting new mechanism to address complex classification problems, which are common in medical or biological information processing applications.

Bo-Suk Yang *et. al.*, [36] have presented a novel scheme to detect faulty products at semi-product stage in an automatic mass product line of reciprocating compressors for small-type refrigerators used in family electrical appliances. They presented the classification accuracy using the ANNs, SVM, LVQ, SOFM and SOFM with LVQ (SOFM-LVQ) and found SOFM-LVQ gives high accuracy and are the best techniques for classifying healthy and faulty conditions of small reciprocating compressors. The result shows SOFM with LVQ can improve the classification performance of SOFM but cannot eliminate the classification error, indicated in the concluding remarks.

Rung-Ching Chen [37] has proposed a web page classification method for extraction of feature vectors from both the LSA and WPFS methods by using a SVM based on a weighted voting schema. The LSA classifies semantically related web pages, offering users more complete information. The experimental results show that the anova kernel function yields the best result of these four kernel functions.

The LSA-SVM, BPN and WVSVM were then compared and demonstrated that the WVSVM yields better accuracy even with a small data set.

Shu-Xin Du *et. al.*, [38] have developed a Weighted support vector machines for classification where penalty of misclassification for each training sample is different. Two weighted support vector machines, namely weighted C-SVM and V-SVM, have been developed for experimenting on breast cancer diagnosis which shows the effectiveness of the proposed methods. They have indicated that, the improvement obtained at the cost of the possible decrease of classification accuracy for the class with large training size and the possible decrease of the total classification accuracy.

Chih-Fong Tsai [39] has presented a two-level stacked generalization scheme composed of three generalizers having color texture of support vector machines (SVMs) for image classification. He has mainly investigated two training strategies based on two-fold cross-validation and non-cross-validation for the proposed classification scheme by evaluating their classification performances, margin of the hyperplane and numbers of support vectors of SVMs. The results show that the non-cross-validation training method performs better, having higher correct classification rates, larger margin of the hyperplane and smaller numbers of support vectors.

Chin-Teng Lin *et. al.*, [40] have proposed a support-vector-based fuzzy neural network (SVFNN) to minimize the training and testing error for better performance. They have developed a learning algorithm consisting of three learning phases is to construct the SVFNN in which the fuzzy rules and membership functions are automatically determined by the clustering principle. To investigate the effectiveness of the proposed SVFNN classification, they applied the corresponding model to various datasets from the UCI Repository and Statlog collection. Experimental results show that the proposed SVFNN for pattern classification can achieve good classification performance with drastically reduced number of fuzzy kernel functions.

Kemal Polat [41] has developed a medical decision making system based on Least Square Support Vector Machine (LSSVM) which was applied on the task of diagnosing breast cancer and the most accurate learning methods was evaluated. He conducted the experiment on the WBCD dataset to diagnose breast cancer in a fully automatic manner using LSSVM. The results strongly suggest that LSSVM can aid in the diagnosis of breast cancer. In his conclusion he has claimed that on the exploration of large data sets the accuracy level may increase.

Sandeep Chaplot *et. al.*, [42] have proposed and implemented a novel approach for classification of MR brain images using wavelet as an input to self-organizing maps and support vector machine .They have noticed classification percentage of more than 94% in case of self organizing maps and 98% in case of support vector machine. They have applied the method only to axial T2-weighted images at a particular depth inside the brain. The same method can be employed for T1-weighted, T2-weighted, proton density and other types of MR images. Also they claim that with the help of above approaches, one can develop software for a diagnostic system for the detection of brain disorders like Alzheimer's, Huntington's, Parkinson's diseases *etc*.

Jin-Hyuk Hong *et. al.*, [43] proposed a novel fingerprint classification method which effectively integrates NBs and OVA SVMs, which produces better accuracy than previously reported in the literature contained in the NIST-4 database. In their proposed method, several popular fingerprint features such as singularities, pseudo codes and the Finger Code were used, and the combination of methods described in the experimental analysis produced better results (90.8% for the five-class classification problem and 94.9% for the four-class classification problem with

1.8% rejection during the feature extraction phase of the Finger Code) than any of the component classifiers.

Fabien Lauer *et. al.*, [44] have proposed different formulations of the optimization problem along with support vector machines (SVMs) for classification task. They have exposed the utility of concerns on the incorporation of prior knowledge into SVMs in their review of the literature. The methods are classified with respect to the categorization into three categories depending on the implementation approach via samples, in the kernel or in the problem formulation. They considered two main types of prior knowledge that can be included by these methods like class invariance and knowledge on the data.

M. Arun Kumar *et. al.*, [45] have enhanced TSVM to least squares TSVM (LSTSVM), which is an immensely simple algorithm for generating linear/nonlinear binary classifiers using two non-parallel hyper planes/ hyper surfaces. In LSTSVM, they have solved the two primal problems of TSVM using proximal SVM (PSVM) idea instead of two dual problems usually solved in TSVM. They have further investigated the application of linear LSTSVM to text categorization using three benchmark text categorization datasets: reuters-21578, ohsumed and 20 Newsgroups (20NG) and based on the Comparison of experimental results, against linear PSVM shows that linear LSTSVM has better generalization on all the three text corpuses considered. Thus they claim that the performance of LSTSVM and PSVM on text categorization can greatly be improved by using it.

Wen Zhang *et. al.*, [46] have implemented the multi-word extraction based on the syntactical structure of the noun multi-word phrases. In order to use the multi-words for representation, they have developed two strategies based on the different semantic level of the multi-words: the first is the decomposition strategy using general concepts for representation and the second is combination strategy using subtopics of the general concepts for representation. IG method was employed as a scale to remove the multi-word from the feature set to study the robustness of the classification performance. Finally, a series of text classification tasks were carried out with SVM in linear and non-linear kernels, respectively, to analyze the effect of different kernel functions on classification performance.

Urmil B. Parikh *et. al.*, [47] have proposed a new SVM based fault classification algorithm for a series compensated transmission line , which uses samples of three phase currents as well as the zero sequence current as input features to the SVMs for identification of the faulted phase(s). They have tested feasibility of the developed technique on an extensive data set of 25,200 test cases covering a wide range of operating conditions and they claim that accuracy of the proposed classification technique has been found to be at least 98%.

Alice Este *et. al.*, [48] have introduced a new classification technique based on Support Vector Machines which is based on a flow representation that expresses the statistical properties of an application protocol. The classification mechanism presents a relatively high complexity during the training phase, especially due to the tuning process of the involved configuration parameters. They have applied the proposed technique to three different data sets and almost in all cases, they found the accuracy of the classifier is very good with classification results (True Positives) going over the 90% mark and in general low False Positive rates.

Chih-Hung Wu *et al.*, [49] have proposed HGASVM which can help innovators or firms in identifying (classifying) and searching critical documents that can assist their strategic decision making process. The contribution of their study is that the proposed algorithm is an effective patent classification system that can ensure the continuous and systematic use of patent information in a company's decision-making processes. By using the HGA-SVM optimization approach, they have significantly improved the necessary steps, computation time and the times of trial-and-error for building an effective SVM classification system.

Takuya Kitamura *et. al.*, [50] have proposed two types of subspace based SVMs (SS_SVMs): subspace-based least squares SVMs (SSLS_SVMs) and subspace-based linear programming SVMs (SSLP_SVMs) where the similarity measure for each class is assumed as the separating hyperplane that separates the associated class with the remaining classes. Also the margin between classes is maximized under the constraints that the similarity measure associated with the class to which a data sample belongs is the largest among all the similarity measures which leads to a linear all-at-once SVM.

Arindam Chaudhuri *et. al.*, [51] have implemented a novel Soft Computing tool viz., FSVM to study the problem of bankruptcy prediction in corporate organizations. The performance of FSVM is illustrated by experimental results which show that they are better capable of extracting useful information from the corporate data than traditional bankruptcy prediction methods. The procedure is easy to implement and is suitable for estimating unique default probability for each firm. The rating estimation done by FSVM is transparent and does not depend on heuristics or expert judgments which imply objectivity and high degree of robustness against user changes.

Jian Qu *et. al.*, [52] have proposed an algorithm which adopts a developed data cleaning algorithm which is based on random sub-sampling validation and support vector classification to discover potential outliers and determines final outliers based on the measure of misclassification rate and uses the well-known sequential backward selection method to identify irrelevant features. The proposed data processing algorithm is applied to the slurry pump system in which the wear degrees of pump impellers are classified. The results indicate that the proposed data processing algorithm is able to achieve effective classifications and suggested that, it is promising to conduct data cleaning before classifications for a better result.

Nahla Barakat *et. al.*, [53] have reviewed on a historical perspective for the SVM area of research and conceptually groups and analyzes the various techniques. In particular, they have proposed two alternative groupings; the first is based on the SVM (model) components utilized for rule extraction, while the second is based on the rule extraction approach. The analysis is followed by a comparative evaluation of the algorithms' salient features and relative performance as measured by a number of metrics. The paper concludes by highlighting potential research directions such as the need for rule extraction methods in the case of SVM incremental and active learning and other application domains, where special types of SVMs are utilized.

Saibal Dutta *et. al.*, [54] have made an attempt to develop a robust heart beat recognition algorithm that can automatically classify normal/ PVC/other heart beats. The work proposes cross-correlation as a formidable feature extraction tool, which when coupled with the LS-SVM classifiers, can be efficiently employed as an automated ECG beat classifier. The performance of the proposed scheme has been evaluated by considering several benchmark signals available in MIT/BIH arrhythmia database and the overall performance was found to be as encouraging as very close to 96%.

Hakan Cevikalp [55] have proposed two new clustering algorithms for the partition of data samples for SVM based BHDTs. The proposed methods have two major advantages over the traditional clustering algorithms like they are suitable when SVMs are used as the base classifier and the most commonly employed k-means clustering algorithm may not be compatible with the SVM classifier as demonstrated in the synthetic database experiments which shows the proposed class based NCuts clustering method seemed more efficient than the proposed sample based clustering method.

Daoliang Li *et. al.*, [56] have developed MSVMs which can classify the foreign fibers in cotton lint in an accurate and quick fashion. Three types of MSVMs, *i.e.*, OAA-DTB MSVM, OAO-VB MSVM and OAO-DAG MSVM were tested with the extracted feature vectors using leave-one-out cross validation. The results indicate that, the OAA-DTB MSVM cannot fulfill the requirement of accuracy of online measurement of content of foreign fiber in cotton lint. However, both of the two one-against-one MSVMs can satisfy the classification accuracy requirement and the OAO-DAG MSVM is the fastest.

John Shawe-Taylor *et. al.*, [57] have presented a review of optimization techniques used for training SVMs. They have shown how to instantiate the KKT conditions for SVMs. Along with the introduction of the SVM algorithms, the characterization of effective kernels has also been presented, which is helpful to understand the SVMs with nonlinear classifiers. For the optimization methodologies applied to SVMs, they have reviewed interior point algorithms, chunking and SMO, coordinate descent, active set methods and Newton's method for solving the primal, stochastic sub gradient with projection and cutting plane algorithms. They believe that the optimization techniques introduced in this paper can be applied to other SVM-related research as well.

Yuan-Hai Shao *et. al.*, [58] have proposed a fast TWSVM-type algorithm-the coordinate descent margin-based twin support vector machine (CDMTSVM) to binary classification. At the same time, they have also proposed a novel coordinate descent method to solve the dual problems. Compared to the original TWSVM, their CDMTSVM is not only faster, but also needs less memory storage. They claim that the obtained computational results on UCI and NDC datasets demonstrate that CDMTSVM obtains classification accuracy better than TWSVM with reduced computational effort for both linear and nonlinear kernels.

Ahmad Kazem *et. al.*, [59] have developed a novel hybrid model based on a chaotic firefly algorithm and support vector regression for stock market price forecasting. Their contribution of the proposed algorithm is mainly the integration of chaotic motion with a firefly algorithm as a simple and novel optimization method. Compared with genetic algorithm-based SVR (SVR-GA), chaotic genetic algorithm-based SVR (SVR-CGA), firefly-based SVR (SVR-FA), artificial neural networks (ANNs) and adaptive neuro-fuzzy inference systems (ANFIS), the proposed model performs best based on two error measures, namely mean squared error (MSE) and mean absolute percent error (MAPE).

E.A. Zanaty [60] have constructed SVMs and computed its accuracy in data classification. They held a comparison between the SVMs and MLP classifier by considering different sizes of data sets with different attributes. Then, they have compared the results of the SVM algorithm to MLP classifier. The proposed GRPF kernel has achieved the best accuracy, especially with the datasets with many attributes. They believe that it greatly reduces the number of operations in the learning mode. It is well seen for large data sets, where SVM algorithm is usually much quicker.

Chin Heng Wan *et. al.*, [61] have implemented a new text document classifier by integrating the K-nearest neighbor (KNN) classification approach with the support vector machine (SVM) training algorithm. The proposed Nearest Neighbor-Support Vector Machine hybrid classification approach is coined as SVM-NN which avoids a major problem of the KNN in determining the appropriate value for parameter K in order to guarantee high classification effectiveness. By considering several benchmark text datasets for their experiments, it is shown that the classification accuracy of the SVM-NN approach has low impact on the value of parameter, as compared to the conventional KNN classification model.

Chien-Shun Lo *et. al.*, [62] have proposed a SVM-based classifiers for MR image classification by presenting two sets of experiments: one set consists of computer-generated phantom images and the other set uses real MR images. From the experimental results, they found the correct rate of SVM classification is significantly better than CM in the case of SNR = 5 dB. Accordingly, they have shown that the SVM has the capability for multi-spectral MR image segmentation and robustness against noise.

Yingjie Tian *et. al.*, [63] have proposed a novel least squares support vector machine, named ε-least squares support vector machine (ε-LSSVM), for binary classification. They claim for improved advantages compared with the plain LSSVM by introducing the ε-insensitive loss function instead of the quadratic loss function into LSSVM. Experimental results on several benchmark datasets show the effectiveness of our method in sparseness, balance performance and classification accuracy.

Zhenning Wu *et. al.*, [64] have proposed a PIM-clustering-based FSVM algorithm for classification problems with outliers or noises. The experiments have been conducted on five benchmark datasets to test the generalization performance of the PIM- FSVM algorithm. Their results have shown that the PIM-FSVM algorithm presents more reasonable memberships and is more robust than other methods used in their paper for classification problems with outliers or noises. Second, the computational complexity of the PIM-FSVM algorithm is presented, which is not more complex or even less complex than other methods.

Zhiquan Qi *et. al.*, [65] have proposed a new Structural Twin Support Vector Machine (called S-TWSVM), which is sensitive to the structure of the data distribution. They firstly pointed out the shortcomings of the existing algorithms based on structural information and designed a new S-TWSVM algorithm and analysis with its advantages and relationships with other algorithms. Theoretical analysis and all experimental results shown that, the S-TWSVM can more fully exploit this prior structural information to improve the classification accuracy.

Himanshu Rai *et. al.*, [66] have introduced a novel and efficient approach for iris feature extraction and recognition. They compared the recognition accuracy with the previous reported approaches for finding better recognition rate than using SVM or Hamming distance alone. They claim for the increase of efficiency, when they used separate feature extraction techniques for SVM and Hamming distance based classifier and proven that the accuracy of the proposed method is excellent for the CASIA as well as for the Chek image database in term of FAR and FRR.

Adil Baykasog lu *et. al.*, [67] have presented a comparative analysis of the performances of GA, DE and FA on both static and dynamic multidimensional knapsack problems. One of the important contributions of their study is the development of FA2, which is designed for a more realistic reflection of the behaviors of fireflies and it requires less computational time. Particularly on stationary environments, FA2 obtains near optimal results with a significantly faster convergence capability. Thus, they assumed that FA2 was found more effective compared to GA, DE and FA for the problems studied.

Jui-Sheng Chou *et. al.*, [68] have proposed several classifiers that can be applied when using CART, QUEST, C5.0, CHAID, and GASVM (a hybrid approach) to predict dispute propensity. In terms of accuracy, GASVM (89.30%) and C5.0 (83.25%) are the two best classification and regression-based models in predicting project disputes. Among all the models, GASVM provides the highest overall performance measurement score (0.871) considering accuracy, precision, sensitivity, and AUC. Notably, with the exception of GASVM, which was developed by the authors and implemented within a mathematical tool, all models are easily executed via open-source or commercial software. Compared to the baseline models (*i.e.*, C5.0, CHAID, CART, and QUEST) and previous work, GASVM provides 5.89–12.95% higher classification accuracy.

Zuriani Mustaffa *et. al.*, [69], have reported empirical results that examine the feasibility of eABC-LSSVM in predicting prices of the time series of interest. The performance of their proposed prediction model was evaluated using four statistical metric, namely MAPE, PA, SMAPE and RMSPE and experimented using three different set of data arrangement, in order to choose the best data arrangement for generalization purposes. In addition, the proposed technique also has proven its capability in avoiding premature convergence that finally leads to a good generalization performance.

Youngdae Kim *et. al.*, [70] have proposed an exact indexing solution for the SVM function queries, which is to find top-k results without evaluating the entire database. They first proposed key geometric properties of the kernel space – ranking instability and ordering stability – which is crucial for building indices in the kernel space. Based on them, they developed an index structure iKernel and processing algorithms and then presented clustering techniques in the kernel space to enhance the pruning effectiveness of the index. According to their experiments, iKernel is highly effective overall producing 1–5% of evaluation ratio on large data sets.

Zhen Yang *et. al.*, [71] have proposed a DE-SVC for pattern recognition. Their research results shown that support vector machines and differential evolution can be used as effective and efficient data processing methods for two-stage HCNs synthesized by fine core/shell particles of PMMA/PAN. Compared with genetic algorithm, average iteration steps and training time of the prediction model based on improved DE-SVC have significantly shortened. They have claimed that SVC has more adaptive learning ability and higher prediction accuracy.

A.D. Dileep *et. al.*, [72] have proposed two novel methods to build a better discriminatory IMK-based SVM classifier by considering a set of virtual feature vectors specific to each class depending on the approaches to multiclass classification using SVMs. They proposed a class-wise IMK based SVM for every class by using components of GMM built for a class and a pair wise IMK based SVM for every pair of classes by using components of GMM built for a pair of classes as the set of virtual feature vectors for that pair of classes in the one-against-one approach to multiclass classification. The performance of the SVM-based classifiers using the proposed class-specific IMKs is studied for speech emotion recognition and speaker identification tasks and compared with that of the SVM-based classifiers using the state-of-the-art dynamic kernels.

Liu Hui [73] have proposed a DFSVM algorithm for classification and adopted for detection cirrhosis from normal hepatic tissue MR imaging. They have extracted Six GLDM based texture features from medical MRI. The Experimental results shown that DFSVM can select important features and strengthen the specific feature by duplication caused by sampling with replacement in iteration. Their proposed DFSVM is compared with typical feature reduction approaches such as PCA, LDA and Weight-Inform Grain, and also compared with typical classifier ANN. The experiment result shown that DFSVM gets both high sensitivity and high specificity.

Yashar Maali *et. al.*, [74] have proposed a self-advising SVM method for the improvement of the SVM performance by transferring more information from the training phase to the testing phase. This information is generated by using misclassified data in the training phase. Experimental results in their study shown improvement in accuracy, and the F-score and statistical tests reveal the significance of these improvements. They claimed that, by using the misclassified data in the training phase, overtraining can be avoided in their proposed method.

Shifei Ding *et. al.*, [75] have enhanced LSPTSVM to nonlinear LSPTSVM(NLSPTSVM)for solving nonlinear classification problems efficiently. Similar to LSPTSVM, NLSPTSVM requires just the solution of two systems of linear equations in contrast to PTSVM which requires solving two QPPs. In order to boost performance of NLSPTSVM, they proposed a novel nonlinear recursive algorithm to improve its performance. Experimental results on synthetic two-moon dataset, several UCI datasets and NDC datasets shown that NLSPTSVM has good nonlinear classification capability.

Zhong Yin [76] have proposed two psycho physiological-data-driven classification frameworks for operator functional states (OFS) assessment in safety-critical human-machine systems with stable generalization ability. They combined the recursive feature elimination (RFE) and least square support vector machine (LSSVM) and used for binary and multiclass feature selection. Feature selection results have revealed that different dimensions of OFS can be characterized by specific set of psycho physiological features. Performance comparison studies shown that reasonable high and stable classification accuracy of both classification frameworks can be achieved if the RFE procedure is properly implemented and utilized.

4. Analytical Discussions, Limitations & Suggestions

This literature review surveys the applications of SVM in diversified fields in connection with the author's background, the application interest and expertise knowledge in the particular field. Some authors have been repeated for different applications. The paper discusses the SVM method applied in a mixture of application areas including medical, engineering, pattern classifications, nuclear component classification, classification problems, prediction, science and other applications, which were extracted from the databases like Elsevier, IEEE X-plore, Springer Link, Taylor Francis and Inderscience.

As the popularity of SVM is increasing day by day, so different research applications relying on SVM must be published, to facilitate the wide broaden scope of SVM, in the academic and practical fields. However, many researchers have pointed some limitations of SVM on which work must be carried out like: (1) The selection of kernel for a problem (2) The functional speed of the machine in training and testing, (3) Slower Convergence rate at testing phase,(4) Choosing good quality kernel parameters ,(5) Large requirements of memory space to implement the model (6) Choosing either parametric or non-parametric method for implementation. This integration of methodologies and cross-disciplinary research may put forward new insights for problem solving with SVM. This paper reviews a no. of major applications using SVM, but still inclusion of some other areas like social, statistical, and behavioral science *etc.*, are needed. Also, the qualitative and quantitative aspects of SVM technique are to be included in our future work.

As shown in Table 1, support vector machine have been applied in almost every application domain including classification, prediction/forecasting, image analysis, pattern recognition, rule extraction and optimization problems. The majority of the applications that we have reviewed, about 54% are in the area of classification. Works in the areas of clustering and forecasting account for 9% and 13%, respectively and others represent 24% each of the reviewed work (Figure 3). The majority of the work review used the MATLAB software for implementing their models.

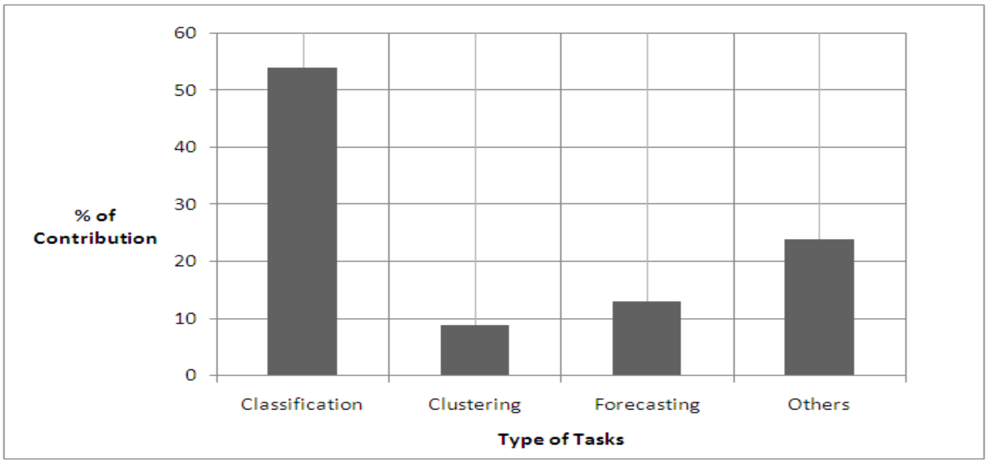

Figure 3. Application of SVM in Data Mining Tasks

Table 1. Summary of Various SVM Models used in Various Applications

Reference	Model Name	Type of Basis Function used	Application Area
[31]	SVM	Gaussian	Drug Design
[32]	SVM	Gaussian	Gene Classification
[33]	SVM	Polynomial	Classification
[34]	SVM	Gradient Descent	Classification
[35]	GSVM	Polynomial	Classification
[36]	SVM	Gaussian	Classification
[37]	WVSVM	Polynomial	Classification
[38]	WC-SVM	Polynomial	Classification
[39]	SVM	Polynomial	Image Classification
[40]	SVFNN	Polynomial	Pattern Classification
[41]	LSSVM	Polynomial	Medical Diagnosis
[42]	SVM	Polynomial	Image Classification
[77]	SVM	Gaussian	Nuclear Component Classification
[43]	OVA SVM	Polynomial	Fingerprint Classification
[44]	SVM	Linear, Polynomial, Gaussian	Classification
[78]	FSVM	Linear, Polynomial, Gaussian	Product Classification
[45]	TSVM	Linear	Pattern Classification
[46]	SVM	Linear and Polynomial	Text Classification
[47]	SVM	Linear, Polynomial, RBF, Sigmoid	Fault Classification
[48]	SVM	Gaussian	TCP Traffic Classification
[49]	HGA-SVM	Linear, Polynomial, Gaussian(RBF)	Patent Classification
[50]	SS-SVM	Polynomial	Pattern Classification
[51]	FSVM	Polynomial	Bankruptcy Prediction
[52]	SVM	Linear	Wear degree Classification
[53]	SVM	Polynomial	Rule Extraction
[54]	LS-SVM	Gaussian	ECG Beat classification
[79]	SVM	Gaussian	Power System Classification
[57]	SVM	Polynomial	Optimization
[58]	CDMTSVM	Quadratic	Classification
[59]	SVR-GA/CGA/FA	RBF	Forecasting
[60]	SVM	GRPF	Classification
[61]	SVM-NN	Polynomial	Text Classification
[80]	F^2SVM	Sinusoid	Voice Classification
[63]	ε-LSSVM	Linear, RBF	Pattern classification
[64]	FSVM	Gaussian	Clustering
[65]	S-TWSVM	Polynomial	Classification
[66]	SVM	Polynomial	Iris Recognition
[68]	GASVM	Polynomial, RBF, Sigmoid	Dispute Classification
[81]	SVM	Cauchy	Risk Minimization
[82]	HS-KCNN-SVM	Polynomial	Text Classification
[83]	SVM	Wavelet	System Identification
[84]	SVM	Gaussian, Polynomial	Fault Diagnosis
[85]	SVR	Gaussian	Reliability Analysis
[86]	SVM	Polynomial	Prediction
[87]	LSSVM	Least Square	Forecasting
[88]	PSVM	Polynomial	Fault Classification
[89]	SVM	Polynomial	Signal Analysis

5. Concluding Remarks

Support Vector Machine is a rapidly increasing field with promise for greater applicability in all domain of research. In this paper the review of different applications of support vector machine is being given focusing on data mining tasks from the year 2001 to 2014. Although, this paper is by no means a meticulous review of the literature in the application of support vector machine to application of data mining, we hope that we have given a passable overview of what is currently happening in this evolving and dynamic area of research. Also the fundamentals of the support vector machines (SVMs) have been discussed along with the different formulations of the optimization problem resulting from the training of such machines. A review of the literature concerning the amalgamation of prior knowledge into SVMs has been exposed. In the last section, a detailed survey statistics report on various papers related to SVM applications published in the standard journals of IEEE, Elsevier, Inderscience, Springer, Taylor Francis are presented. After analyzing the literature survey, the following are some of the issues recognized that can be taken further to do the research. Even though various techniques have been used in the literature survey, still there is a need of best techniques to solve the following research issues.

- The Learning phase of SVM scale with the number of training data points [90].
- If the data set size increases, the learning phase can lead to a slower process.
- Some activities like Data cleaning, Data Transformation and Outlier detection are important problem for any type of data sets since some of the attributes values cannot be obtained usually. So handling of missing values for classification or forecasting problem is a challenging task.
- Complexity of handling the large dataset for any application is a recent issue since most of classification algorithms are not suitable to handle it.
- Selection of most apposite sample of data for classification instead of the intact data is another risk for getting better result.

Selecting the suitable classification techniques without much computation complexity is another positive direction but the effectiveness should not be affected.

Acknowledgements

This work is supported by Department of Science & Technology (DST), Ministry of Science & Technology, New Delhi, Govt. of India, under grants No. DST/INSPIRE Fellowship/2013/585.

References

[1] H. Gish, "A probabilistic approach to the understanding and training of neural network classifiers", IEEE international conference acoustic, speech signal process, vol. 3, **(1990)**, pp. 1361–1364.

[2] G. P. Zhang, "Neural networks for classification: a survey", IEEE Trans Syst Man Cybern C, vol. 30, vol. 4, **(2000)**, pp. 451-462.

[3] D. Michie, D. J. Spiegel Halter and C. C. Taylor, "Machine learning, neural and statistical classification", Ellis Horwood, NewYork, **(1994)**.

[4] M. Adya and F. Collopy, "How effective are neural networks at forecasting and prediction? A review and evaluation", Forecast, vol. 17, **(1998)**, pp. 481-495.

[5] J. L. Callen, C. C. Y. Kwan, P. C. Y. Yip and Y. Yuan, "Neural network forecasting of quarterly accounting earnings", Forecast, vol. 12, **(1996)**, pp. 475-482.

[6] K. B. Church and S. P. Curram, "Forecasting consumers expenditure: a comparison between econometric and neural network models", Forecast, vol. 12, **(1996)**, pp. 255–267.

[7] J. T. Connor, R. D. Martin and L. E. Atlas, “Recurrent neural networks and robust time series prediction”, IEEE Trans Neural Netw., vol. 51, no. 2, **(1994)**, pp. 240–254.
[8] M. Cottrell, B. Girard, Y. Girard, M. Mangeas and C. Muller, “Neural modeling for time series: a statistical stepwise method for weight elimination”, IEEE Trans Neural Netw., vol. 6, no. 6, **(1995)**, pp. 1355–1364.
[9] J. J. Faraway and C. Chatfield, “Time series forecasting with neural networks: a comparative study using the airline data”, Appl Stat., vol. 47, **(1998)**, pp. 231–250.
[10] D. Fletcher and E. Goss, “Forecasting with neural networks— an application using bankruptcy data”, Inf Manag, vol. 24, **(1993)**, pp. 159–167.
[11] W. L. Gorr, “Research prospective on neural network forecasting”, Forecast, vol. 10, **(1994)**, pp. 1–4.
[12] H. S. Hippert, C. E. Pedreira and R. C. Souza, “Neural networks for short-term forecasting: a review and evaluation”, IEEE Trans Power Syst., vol. 16, no. 1, **(2001)**, pp. 44-55.
[13] M. R. Belli, M. Conti, P. Crippa and C. Turchetti, “Artificial neural networks as approximators of stochastic processes”, Neural Netw., vol. 12, no. 4-5, **(1999)**, pp. 647–658.
[14] J. L. Castro, C. J. Mantas and J. M. Benitez, “Neural networks with a continuous squashing function in the output are universal approximators”, Neural Netw., vol. 13, no. 6, **(2000)**, pp. 561–563.
[15] K. Funahashi, “On the approximate realization of continuous mappings by neural networks”, Neural Netw., vol. 2, **(1989)**, pp. 183-192.
[16] R. Andrews, J. Diederich and A. B. Tickle, “Survey and critique of techniques for extracting rules from trained artificial neural networks”, Knowl Based Syst., vol. 8, no. 6, **(1995)**, pp. 373–389.
[17] J. L. Castro, I. Requena and J. M. Benitez, “Interpretation of artificial neural networks by means of fuzzy rules”, IEEE Trans Neural Netw., vol. 13, no. 1, **(2002)**, pp. 101-116.
[18] R. Setiono, W. K. Leow and J. Zurada, “Extraction of rules from artificial neural networks for nonlinear regression”, IEEE Trans Neural Network, vol. 13, no. 3, **(2002)**, pp. 564–577.
[19] R. Setiono and J. Y. L. Thong, “An approach to generate rules from neural networks for regression problems”, Eur J Oper Res., vol. 155, **(2004)**, pp. 239-250.
[20] P. J. G. Lisboa, “A review of evidence of health benefit from artificial neural networks in medical intervention”, Neural Netw., vol. 15, **(2002)**, pp. 11-39.
[21] L. G. Protney and M. P. Watkins, “Foundations of clinical research: applications to practice”, Prentice-Hall, Princeton, **(2000)**.
[22] S. M. Hosseini-Nezhad, T. S. Yamashita, R. A. Bielefeld, S. E. Krug and Y. H. Pao, “A neural network approach for the determination of inter hospital transport mode”, Comput Biomed Res., vol. 28, no. 4, **(1995)**, pp. 319-334.
[23] C. J. Burges, “A tutorial on support vector machines for pattern recognition”, Data Mining and Knowledge Discovery, vol. 2, **(1998)**, pp. 1-43.
[24] V. Vapnik, “Statistical Learning Theory”, John Wiley and Sons, **(1998)**.
[25] E. Sami, “Support Vector Machines for classification and locating faults on transmission lines”, Applied Soft Computing, vol. 12, **(2012)**, pp. 1650–1658.
[26] T. Joachims, “Learning to classify text using support vector machines”, Kluwer Academic Publishers, **(2002)**.
[27] S. N. Sivanandam, S. Sumathi and S. N. Deepa, “Introduction to Neural Networks using MATLAB 6.0”, Tata McGraw Hill Education Pvt. Ltd., **(2006)**.
[28] J. Li, N. Allinson, D. Tao and X. Li, “Multitraining support vector machine for image retrieval”, IEEE Transaction on Image Processing, vol. 15, **(2006)**, pp. 3597–3601.
[29] V. N. Vapnik, “Estimation of Dependences Based on Empirical Data”. Springer-Verlag, NY., **(1982)**.
[30] M. Pontil and A. Verri, “Support vector machines for 3D object recognition”, IEEE Transaction on PAMI, vol. 20, no. 6, **(1998)**, pp. 637–646.
[31] R. Burbidge, M. Trotter, B. Buxton and S. Holden, “Drug design by machine learning: support vector machines for pharmaceutical data analysis”, Computers and Chemistry, vol. 26, **(2001)**, pp. 5-14.
[32] G. Valentini, “Gene expression data analysis of human lymphoma using support vector machines and output coding ensembles”, Artificial Intelligence in Medicine, vol. 26, **(2002)**, pp. 281–304.
[33] L. Shutao, J. T. Kwok, H. Zhua and Y. Wang, “Texture classication using the support vector machines”, Pattern Recognition, vol. 36, **(2003)**, pp. 2883-2893.
[34] Y. Zhan and D. Shen, “Design efficient support vector machine for fast classification”, Pattern Recognition, vol. 38, **(2005)**, pp. 157-161.
[35] Y. Tang, B. Jin, Y. Sun and Y. Zhang, “Granular support vector machines for medical binary classification problems”, Proceedings of 2004 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB '04, **(2004)**, pp. 73-78.
[36] B. Yang, W. Hwang, D. Kim and A. Tan, “Condition classification of small reciprocating compressor for refrigerators using artificial neural networks and support vector machines”, Mechanical Systems and Signal Processing, vol. 19, **(2005)**, pp. 371–390.
[37] R. Chen and C. Hsieh, “Web page classification based on a support vector machine using a weighted vote schema”, Expert Systems with Applications, vol. 31, **(2006)**, pp. 427-435.
[38] S. Du and S. Chen, “Weighted support vector machine for classification”, IEEE International Conference on Systems, Man and Cybernetics, vol. 4, **(2005)**, pp. 3866-3871.

[39] C. Tsai, “Training support vector machines based on stacked generalization for image classification”, Neurocomputing, vol. 64, **(2005)**, pp. 497–503.
[40] C. Lin, C. Yeh, S. Liang, J. Chung and N. Kumar, “Support-vector-based fuzzy neural network for pattern classification”, IEEE Transactions on Fuzzy Systems, vol. 14, no. 1, **(2006)**, pp. 31-41.
[41] K. Polat and S. Gune, “Breast cancer diagnosis using least square support vector machine”, Digital Signal Processing, vol. 17, **(2007)**, pp. 694–701.
[42] S. Chaplot, L. M. Patnaik and N. R. Jagannathan, “Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network”, Biomedical Signal Processing and Control, vol. 1, **(2006)**, pp. 86–92.
[43] J. Hong, J. Min, U. K. Cho and S. B. Cho, “Fingerprint classification using one-vs-all support vector machines dynamically ordered with naïve Bayes classifiers”, Pattern Recognition, vol. 41, **(2008)**, pp. 662-671.
[44] F. Lauer and G. Bloch, “Incorporating prior knowledge in support vector machines for classification: A review”, Neurocomputing, vol. 71, **(2008)**, pp. 1578–1594.
[45] A. M. Kumar and M. Gopal, “Least squares twin support vector machines for pattern classification”, Expert Systems with Applications, vol. 36, **(2009)**, pp. 7535–7543.
[46] W. Zhang, T. Yoshida and X. Tang, “Text classification based on multi-word with support vector machine”, Knowledge-Based Systems, vol. 21, **(2008)**, pp. 879–886.
[47] U. B. Parikh, B. Das and R. Maheshwari, “Fault classification technique for series compensated transmission line using support vector machine”, Electrical Power and Energy Systems, vol. 32, **(2010)**, pp. 629–636.
[48] A. Este, F. Gringoli and L. Salgarelli, “Support Vector Machines for TCP traffic classification”, Computer Networks, vol. 53, **(2009)**, pp. 2476–2490.
[49] C. Wu, Y. Ken and T. Huang, “Patent classification system using a new hybrid genetic algorithm support vector machine”, Applied Soft Computing, vol. 10, **(2010)**, pp. 1164–1177.
[50] T. Kitamura, S. Takeuchi, S. Abe and K. Fukui, “Subspace-based support vector machines for pattern classification”, Neural Networks, vol. 22, **(2009)**, pp. 558-567.
[51] A. Chaudhuri and K. De, “Fuzzy Support Vector Machine for bankruptcy prediction”, Applied Soft Computing, vol. 11, **(2011)**, pp. 2472–2486.
[52] J. Qu and M. J. Zuo, “Support vector machine based data processing algorithm for wear degree classification of slurry pump systems”, Measurement, vol. 43, **(2010)**, pp. 781–791.
[53] N. Barakat and A. P. Bradley, “Rule extraction from support vector machines: A review”, Neurocomputing, vol. 74, **(2010)**, pp. 178–190.
[54] S. Dutta, A. Chatterjee and S. Munshi, “Correlation technique and least square support vector machine combine for frequency domain based ECG beat classification”, Medical Engineering & Physics, vol. 32, **(2010)**, pp. 1161–1169.
[55] H. Cevikalp, “New clustering algorithms for the support vector machine based hierarchical classification”, Pattern Recognition Letters, vol. 31, **(2010)**, pp. 1285–1291.
[56] D. Li, W. Yang and S. Wang, “Classification of foreign fibers in cotton lint using machine vision and multi-class support vector machine”, Computers and Electronics in Agriculture, vol. 74, **(2010)**, pp. 274–279.
[57] J. Taylor and S. Sun, “A review of optimization methodologies in support vector machines”, Neurocomputing, vol. 74, **(2011)**, pp. 3609-3618.
[58] Y. H. Shao and Y. N. Deng, “A coordinate descent margin based-twin support vector machine for classification”, Neural Networks, vol. 25, **(2012)**, pp. 114–121.
[59] A. Kazem, E. Sharifi, F. K. Hussain, M. Saberi and O. K. Hussain, “Support vector regression with chaos-based firefly algorithm for stock market price forecasting”, Applied Soft Computing, vol. 13, **(2013)**, pp. 947–958.
[60] E. A. Zanaty, “Support Vector Machines (SVMs) versus Multilayer Perception (MLP) in data classification”, Egyptian Informatics Journal, vol. 13, **(2012)**, pp. 177–183.
[61] C. H. Wan, L. H. Lee, R. Rajkumar and D. Isa, “A hybrid text classification approach with low dependency on parameter by integrating K-nearest neighbor and support vector machine”, Expert Systems with Applications, vol. 39, **(2012)**, pp. 11880–11888.
[62] C. S. Lo and C. M. Wang, “Support vector machine for breast MR image classification”, Computers and Mathematics with Applications, vol. 64, **(2012)**, pp. 1153–1162.
[63] Y. Tian, X. Ju, Z. Qi and Y. Shi, “Efficient sparse least squares support vector machines for pattern classification”, Computers and Mathematics with Applications, vol. 66, **(2013)**, pp. 1935-1947.
[64] Z. Wu, H. Zhang and J. Liu, “A fuzzy support vector machine algorithm for classification based on a novel PIM fuzzy clustering method”, Neurocomputing, vol. 125, **(2014)**, pp. 119–124.
[65] Z. Qi, Y. Tian and Y. Shi, “Structural twin support vector machine for classification”, Knowledge-Based Systems, vol. 43, **(2013)**, pp. 74–81.
[66] H. Rai and A. Yadav, “Iris recognition using combined support vector machine and Hamming distance approach”, Expert Systems with Applications, vol. 41, **(2014)**, pp. 588–593.

[67] A. Baykasoglu and F. B. Ozsoydan, "An improved firefly algorithm for solving dynamic multidimensional knapsack problems", Expert Systems with Applications, vol. 41, **(2014)**, pp. 3712–3725.
[68] J. S. Chou, M. Y. Cheng, Y. W. Wu and A. D. Pham, "Optimizing parameters of support vector machine using fast messy genetic algorithm for dispute classification", Expert Systems with Applications, vol. 41, **(2014)**, pp. 3955–3964.
[69] Z. Mustaffa, Y. Yusof and S. S. Kamaruddin, "Enhanced artificial bee colony for training least squares supportvector machines in commodity price forecasting", Journal of Computational Science, vol. 5, no. 2, **(2014)** March, pp. 196–205.
[70] Y. Kim, I. Ko, W. S. Han and H. Yu, "iKernel: Exact indexing for support vector machines", Information Sciences, vol. 257, **(2014)**, pp. 32–53.
[71] Z. Yang, Q. Yu, W. Dong, X. Gu, W. Qiao and X. Liang, "Structure control classification and optimization model of hollow carbon nanosphere core polymer particle based on improved differential evolution support vector machine", Applied Mathematical Modelling, vol. 37, **(2013)**, pp. 7442-7451.
[72] A. D. Dileep and S. C. Chandra, "Class-specific GMM based intermediate matching kernel for classification of varying length patterns of long duration speech using support vector machines", Speech Communication, vol. 57, **(2014)**, pp. 126–143.
[73] L. Hui, G. D. Mei and L. Xiang, "Cirrhosis classification based on MRI with duplicative-feature support vector machine (DFSVM)", Biomedical Signal Processing and Control, vol. 8, **(2013)**, pp. 346– 353.
[74] Y. Maali and A. A. Jumaily, "Self-advising support vector machine", Knowledge-Based Systems, vol. 52, **(2013)**, pp. 214–222.
[75] S. Ding and X. Hua, "Recursive least squares projection twin support vector machines for nonlinear classification", Neurocomputing, **(2014)**.
[76] Z. Yin and J. Zhang, "Operator functional state classification using least-square support vector machine basedrecursive feature elimination technique", Computer methods and programs in biomedicine, vol. 113, **(2014)**, pp. 101-115.
[77] S. R. M. Claudio and E. Zio, "A support vector machine integrated system for the classification of operation anomalies in nuclear components and systems", Reliability Engineering and System Safety, vol. 92, **(2007)**, pp. 593-600.
[78] M. D. Shieh and C. C. Yang, "Classification model for product form design using fuzzy support vector machines", Computers & Industrial Engineering, vol. 55, **(2008)**, pp. 150–164.
[79] H. Eris, A. Ucar and A. Demir, "Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines", Electric Power Systems Research, vol. 80, **(2010)**, pp. 743–752.
[80] S. Scherer, J. Kane, C. Gobl and F. Schwenker, "Investigating fuzzy-input fuzzy-output support vector machines for robust voice quality classification", Computer Speech and Language, vol. 27, **(2013)**, pp. 263-287.
[81] L. Suzhen, L. Xiangjie and Y. Gang, "Supervisory predictive control of weighted least square support vector machine based on Cauchy distribution", The 26th Chinese Control and Decision Conference (CCDC), DOI: 10.1109/CCDC.2014.6852789, **(2014)**, pp. 3523-3526.
[82] Y. H. Chen, Y. F. Zheng, J. F. Pan and N. Yang, "A Hybrid Text Classification Method Based on K-Congener-Nearest-Neighbors and Hypersphere Support Vector Machine", International Conference on Information Technology and Applications (ITA), DOI: 10.1109/ITA.2013.120, **(2013)**, pp. 493-497.
[83] Z. Lu, J. Sun and K. Butts, "Multiscale Asymmetric Orthogonal Wavelet Kernel for Linear Programming Support Vector Learning and Nonlinear Dynamic Systems Identification", IEEE Transactions on Cybernetics, DOI: 10.1109/TCYB.2013.2279834, vol. 44, no. 5, **(2014)**, pp. 712-724.
[84] Y. Zhang, X. Ma, Y. Zhang and J. Yang, "Support vector machine of the coal mine machinery equipment fault diagnosis", IEEE International Conference on Information and Automation (ICIA), 1141–1146, DOI: 10.1109/ICInfA.2013.6720467, **(2013)**.
[85] S. R. Dindarloo, "Support vector machine regression analysis of LHD failures", International Journal of Mining, Reclamation and Environment, DOI:10.1080/17480930.2014.973637, **(2014)**.
[86] D. Kim, "Prediction performance of support vector machines on input vector normalization methods", International Journal of Computer Mathematics, DOI: 10.1080/00207160410001684325, **(2010)**.
[87] A. Shabri and Suhartono, "Stream flow forecasting using least-squares support vector machines", Hydrological Sciences Journal, DOI: 10.1080/02626667.2012.714468, **(2012)**.
[88] N. R. Sakthivel, V. Sugumaran and B. B. Nair, "Application of Support Vector Machine (SVM) and Proximal Support Vector Machine (PSVM) for fault classification of monoblock centrifugal pump", Int. J. of Data Analysis Techniques and Strategies, vol. 2, no. 1, **(2010)**, pp. 38-61.
[89] M. Saimurugan and K. I. Ramachandran, "A comparative study of sound and vibration signals in detection of rotating machine faults using support vector machine and independent component analysis", Int. J. of Data Analysis Techniques and Strategies, vol. 6, no. 2, **(2014)**, pp. 188-204.
[90] R. Strack, V. Kecman, B. Strack and Q. Li, "Sphere Support Vector Machines for large classification tasks", Neurocomputing, vol. 101, **(2013)**, pp. 59–67.

Authors

Janmenjoy Nayak, is a Research Fellow in the Department of Computer Sc. Engineering & Information Technology, Veer Surendra Sai University of Technology (VSSUT), Burla, Odisha, India. He has been awarded INSPIRE fellowship of Department of Science and Technology, Ministry of Science and Technology, Govt. of India for pursuing Ph.D. in Computer Science & Engineering at VSSUT, Burla, Sambalpur, Odisha, India. His area of interest includes Data Mining, Artificial Intelligence, Pattern Recognition and Soft Computing.

Bighnaraj Naik, is working as an Assistant Professor in the Department of Master in Computer Application, Veer Surendra Sai University of Technology (VSSUT), Burla, Odisha, India. Currently, he is a Ph.D. Research Scholar in the Department of Computer Sc. Engineering & Information Technology, VSSUT, Burla, Odisha, India. His area of interest includes Data Mining, Machine Learning and Soft Computing.

Dr. Himansu Sekhar Behera, is working as a Reader in the Department of Computer Sc. Engineering & Information Technology, Veer Surendra Sai University of Technology (VSSUT), Burla, Odisha. His research interest includes Data Mining and Intelligent Computing.

