An Improved Ant Colony Optimization for Parameter Optimization using Support Vector Machine

Srujana Rongali, Radhika Yalavarthi

Abstract. Support Vector Machine (SVM) is one of the significant classification technique and it can be applied in various areas like meteorology, financial data analysis etc. The performance of SVM is influenced by parameters like C, which is cost constant and kernel parameter. In this paper, an improved Ant Colony Optimization (IACO) technique is proposed to optimize the parameters of SVM. To evaluate the proposed approach, the experiment adopts five benchmark datasets. The developed approach was compared with the ACO-SVM algorithm proposed by Zhang et al. The experimental results of the simulation show that performance of the proposed method is encouraging.

Keywords: Support vector machines, Ant colony optimization, Parameter optimization

I. INTRODUCTION

Support vector machines (SVMs) are administered learning approaches and were initially proposed by Vapnik [1] in the mid 90's. SVM is based on structural risk minimization principle [2]. SVMs can be used to solve several problems such as Fault Detection [2,3], Fault Prediction [4], Forecasting [5,6], Aquaculture water quality prediction [7], Predicting Bankruptcy [8], Predicting wave transmission [9], Fault Diagnosis [10] etc.

The training data is mapped to a high dimensional feature space [11] through a mapping function, thereby constructing a separating hyper plane that maximizes the margin. SVM can efficiently be used for building classification model using the training data and the model is tested using test data [11]. Proper parameter setting improves the SVM classification accuracy [12]. The parameters to be optimized are C and γ [12].

In the recent past researchers have suggested several methods to select the parameters of the classification model. Huang et al. [12] employed grid algorithm and genetic algorithm based approach for parameter searching on 11 real-world datasets [12][12] from the UCI database [12]. Huang et al. compared their proposed genetic algorithm based approach with grid algorithm and observed that genetic algorithm based approach significantly improved the classification accuracy. Aydin et al. [13] suggested a multi- objective artificial immune algorithm to optimize the SVM parameters and used the algorithm for diagnosing faults of induction motors. Zhao et al. [14] presented a new intelligent back analysis method for recognizing the geomechanical parameters through combining particle swarm optimization, support vector machines and numerical analysis. Eitrich et al. [15] proposed a framework on an automated tuning of the learning parameters of SVM. Derivative-free optimization methods such as APPSPACK is used for this task. Hou et al. [4] proposed a short-term fault prediction method using SVMs, an evolutionary optimization strategy with covariance matrix adaptation (CMA-ES) is used to optimize the parameters of SVM. Samanta et al. [3] presented a genetic algorithm based feature selection method, where the classification model parameters and selection of input features are optimized using this approach. Wang et al. [16] proposed a genetic algorithm based smooth twin parametric-margin support vector machine (STPMSVM) [16], which selects the parameters efficiently in addition to discriminative feature selection [16]. Zhao et al. employed genetic algorithm with feature chromosomes to optimize the feature subset and SVM parameters [17].

Wu et al. [8] employed a real-valued genetic algorithm to optimize the parameters of SVM for predicting bankruptcy. Reif et al. [18] presented the ideas from meta-learning for classifier parameter optimization to provide initial points to the genetic algorithm, such that acceptable accuracy is obtained. Zhang et al. [19] merged simulated annealing with genetic algorithm (SA-GA), in order to choose the kernel parameter of SVM. The strength of SA overcomes the weakness of GA and vice versa. It has been observed that the performance is better if both SA & GA are combined than applying SA or GA independently for choosing the kernel parameters of SVM. Adankon et al. [20] proposed regularization term to form new objective function for semi-supervised SVM (S^3VM). Genetic algorithm (GA) is applied to S^3VM to optimize the objective function, besides designing the specific genetic operators and certain heuristics to improve the optimization task. The author proved that his proposed S^3VM method resulted in more accurate classification than classical optimization methods used for S^3VM. Boardman et al. [21] proposed a heuristic method for selection of free parameters in support vector machine to enhance generalization performance and selection of training set.

Manuscript published on 28 February 2017.
* Correspondence Author (s)
Srujana Rongali, Department of Computer Science & Engineering, GIT, GITAM University, Visakhapatnam, India. E-mail: rsrujana8@gmail.com
Radhika Yalavarthi, Department of Computer Science & Engineering, GIT, GITAM University, Visakhapatnam, India., E-mail: yradhikacse@gmail.com

Based on the inclusion of a model complexity measure and an intensity-weighted centre of mass, the proposed heuristic method produced comparable results with that of grid search. Genetic algorithm based SVM parameter optimization was used by Yu et al. [22] where the parameters are automatically optimized. Yu et al. compared their proposals with gradient descending method and they observed that their method yields promising results. Huaitie et al. [23] employed genetic simulated annealing algorithm and support vector bound which is modified as criterion function based on hybrid optimization strategy. Authors tested this algorithm with benchmark data sets such as banana, flare-solar, heart, thyroid etc., and compared using cross validation methods and observed that the accuracy of SVM parameter selection and generalization performance of SVM has been improved. Particle swarm optimization (PSO) has been used along with SVM by Lin et al. [24] for optimal parameters search and to obtain a feature subset that has been used to train & test the model. The authors compared the proposed method with grid search method on public UCI datasets like australian, bio-informatics, breast cancer, iris, etc. and they observed that the classification accuracy improved noticeably. ACO algorithm was introduced by Macro Dorigo and team in early 90's for the solution of combinational optimization [25]. ACO deals with discrete optimization problems [26]. ACO involves basic mathematical operations. Due to the characteristics like parallelism and distribution [27], ACO processes massive data. Researches has indicated that ACO has the capability of solving global optimization problems. Zhang et al. [25] developed an ant colony optimization (ACO) based technique for optimizing SVM parameters, i.e. ACO-SVM model. This model has been tested on five UCI benchmark datasets namely breast cancer, diabetes, heart, thyroid and titanic [25]. The authors compared the results with grid-based SVM algorithm, 5-fold cross validation, etc. Zhang et al. [25] stated that their model resulted in more accurate results and hence they suggest that ACO algorithm is feasible for optimizing SVM parameters. An enhanced ant colony optimization (IACO) was proposed by Li et al. [2] for selection of SVM parameters. Here meshing is applied to adjust the parameters optimized. Li et al. used rolling bearing vibration signal investigated data and compared with genetic algorithm, cross validation and standard ACO. It has been observed that IACO-SVM performs better in optimizing the parameters of SVM [2]. Alwan et al. [26] employed continuous ant colony optimization (ACO) to optimize parameters of SVM, where there is no need to discretize the continuous value. Therefore loss of information will not happen and classification accuracy will not be affected. The authors implemented the algorithm in seven UCI datasets, such as Australian, pima-indian diabetes, heart etc. The experimental results proved that the proposed method has performed well in terms of computational speed, training time and test accuracy, when compared to grid search technique.

II. SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is one of the emerging data classification technique introduced by Vapnik [1] and is based on structural risk minimization principle [2]. SVM can be used to solve both linear and nonlinear classification problems [28–30]. A classifier model is built by considering an input vectorX_i. The model is used to classify an unknown input vector.

Given a set of training data $T = \{X_i, Y_i\}_{i=1}^{n}$, where X_i, Y_i and n represent the input data, class label and number of samples respectively. The Linear Discriminate Function (LDF) for this input vector can be written as:

$$g(X_i) = W^T.\lambda(X_i) + b \; . \qquad \textbf{(1)}$$

Where $W^T.\lambda(X_i) + b$ is a plane i.e. a straight line in a 2-dimensional space, W is n-dimensional weight vector and b is bias value.

A parameter γ is set in such a way that it will separate the input feature vector and the hyper plane with maximum separation such that any disturbance or noise may not result in erroneous classification. Hence, the linear function can also be written as:

$$g(X) = Y_i(W^T.\lambda(X_i) + \; b) > \gamma \; . \qquad \textbf{(2)}$$

So for designing a support vector machine the value of $W^T.\lambda(X_i) + b$ has to be maximum and can be obtained by lowering the $||W^2||$ and maximizing b in above equation. This can be done by Lagrange function $L(W, b, \alpha)$ [10]

$$L(W, b, \alpha) = \frac{1}{2}||W^2|| - \sum_{i=1}^{n} \alpha_i [Y_i(W^T.\lambda(X_i) + \; b) - 1] \; . \qquad \textbf{(3)}$$

In-order to obtain optimal generalization, slack variable is introduced

$$min\, \lambda(W, Ɣ) = \frac{1}{2}||W^2|| + C \sum_{i=1}^{n} Ɣ_i \; . \qquad \textbf{(4)}$$

Where C is the cost constant. The Lagrange function $L(W, b, \alpha)$ [10] can be written as

$$L(W, b, \alpha, Ɣ) = \frac{1}{2}||W^2|| + C \sum_{i=1}^{n} Ɣ_i - \sum_{i=1}^{n} \alpha_i [Y_i(W^T.\lambda(X_i) + \; b) - 1] \sum_{i=1}^{n} \Pi_i Ɣ_i \; . \qquad \textbf{(5)}$$

Taking derivative of L pertaining to W, b and Ɣ results in

$$w = \sum_{i=1}^{m} \alpha_i Y_i \lambda(X_i) \; . \qquad \textbf{(6)}$$

$$\alpha_i [Y_i(W^T.\lambda(X_i) + \; b) - 1] = 0 \; . \qquad \textbf{(7)}$$

By using Karush–Kuhn–Tucker (KKT) conditions [1], Equation (7) is used to determine b and it can be written as

$$b = \frac{1}{N_s} \sum_{0<\alpha_j<C} Y_j - W^T.\lambda(X_j) \; . \qquad \textbf{(8)}$$

For unknown sample X, the classification function is

$$f(X) = sign(W^T.\lambda(X) + \; b) \; . \qquad \textbf{(9)}$$

Substituting Equations (6) and (8) in Equation (9), nonlinear classification function [1] is obtained as

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

$$f(X) = sign\left(\sum_{i=1}^{N} \alpha_i Y_i K(xi,xj) + \frac{1}{N_s}\sum_{0<\alpha_j<C}\left(Y_i - \sum_{i=1}^{N} \alpha_i Y_i K(xi,xj)\right)\right). \quad \textbf{(10)}$$

Various kernel functions, namely Radial Kernel/ Gaussian, Sigmoid Kernel/ Multilayer Perceptron, Polynomial Kernels etc. can be used.

Usually SVM accuracy is influenced by certain user defined parameters namely C and kernel parameter [31]. Careful selection of SVM parameters affects the accuracy of the SVM model. A comprehensive examination can be performed over the parameter space but it is computationally complex. Trial and error method consumes more time and the obtained results may not be trustworthy [32].

Generalization error estimation and gradient descent [25] are the most elaborated techniques for parameter optimization. In 2002, Chapelle et al [33] proposed a method for minimizing the error along with gradient descent algorithm where the methods are updated iteratively.

Zhang et al. [25] presented an ACO algorithm, to optimize the SVM classifier parameters. Regularization constant C and kernel function parameters have an effect on the ability of SVM. The authors [25] used ant's solution to represent C and sigma, where sigma is the RBF kernel parameter. RBF is used as it can effectively deal with high dimensional data.
The overall process of ACO as mentioned by Zhang et al. [25] is as follows:
Primarily, Initialization of parameters and variables is performed followed by calculation of grid size.
The pheromone levels of the parameter combination are same at all nodes [25], the ants are placed randomly and state transition rule is applied, to construct solution for c and σ.

$$P_{ij} = \frac{\tau_{ij}}{\sum_{i=1}^{N} \tau_{ij}}. \quad \textbf{(11)}$$

P_{ij} = Problem model, τ_{ij} = Pheromone values.

$$Minimize\ T = \frac{1}{l}\sum_{i=1}^{l} \psi\left(-y_i^1 . f(x_i^1)\right). \quad \textbf{(12)}$$

Ψ= Step Function, f=Decision function of SVM, T= Pheromone trail parameters, x_i, y_i =input and output pairs of a training set.

Then state updating rule is applied to update the pheromone level of the finest parameter set.

$$\tau_{ij}^{new} = (1-P).\tau_{ij}^{old} + \frac{Q}{e^T}. \quad \textbf{(13)}$$

τ = Pheromone value, $P = (m_j - \Delta) * h_j$, Q = Pheromone intensity.

Choice of next node depends on pheromone quality, where the pheromone is maximum and the rest of pheromone is reduced.

$$v_j^{lower} \leftarrow v_j^{lower} + (m_j - \Delta) * h_j;\ \Delta = coefficient. \quad \textbf{(14)}$$

$$v_j^{upper} \leftarrow v_j^{upper} + (m_j - \Delta) * h_j. \quad \textbf{(15)}$$

The above process iterates, until the predefined accuracy is greater than the grid interval h_j [19][25]. And finally optimal parameter values are obtained.

Reducing the search space after each iteration increases the overhead and this needs the selection of a pair of C and sigma values which may generate local maxima for the SVM classification. As an ant proceeds from one position to another, there is no assurance that the next node is not occupied. Therefore it leads to high consumption of ants. The next section discusses an improved ACO algorithm that overcomes the disadvantages in the technique proposed by Zhang et al.

III. PROPOSED IMPROVED ACO ALGORITHM

The following steps illustrate the detailed process of the proposed algorithm:

Step 1: In the initial step, an ant establishes a possible best path. Initial values of the cost constant C and RBF kernel parameter γ are fixed. Values of number of ants, iterations, records in test data, grids, ranges for C and γ are also fixed.

Step 2: Then C and γ are meshed according to the grid size as Pi= (Pmax-Pmin)/N, where Pmax and Pmin specify the upper and lower bounds of the parameter, N= Grid Size and Pi=Parameter interval.

Step 3: There is a uniform distribution of the pheromone levels over the grid and the ants are placed stochastically all over the grid.

Step 4: The second phase involves training the SVM with the parameters represented by the ants' locations. An SVM model is generated for each parameter pair and is tested using the test data.

Step 5: The third phase relates to comparing the accuracy of the model associated with each ant with the best accuracy of SVM found till now.

Step 6: If the accuracy of a new SVM model is better than the previous one, update the latest accuracy and record the corresponding C and γ values.

Step 7: The fourth phase involves, updation of pheromone to the node with a parameter combination, to attract the forthcoming ants. Search the vicinity of each ant to identify a node which has a higher pheromone level than the one it is currently occupying. If such node is found, move the ant to that node.

Step 8: If not, move the ant to a random node in its neighbourhood ensuring that the node that an ant moves to is not being occupied by another ant currently.

Step 9: Finally the process continues until the pre-specified numbers of iterations are met. Classification accuracy of the best SVM model, along with its corresponding C and γ values are observed.

The overall process of the improved ACO algorithm is depicted in figure 1.

*Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)*

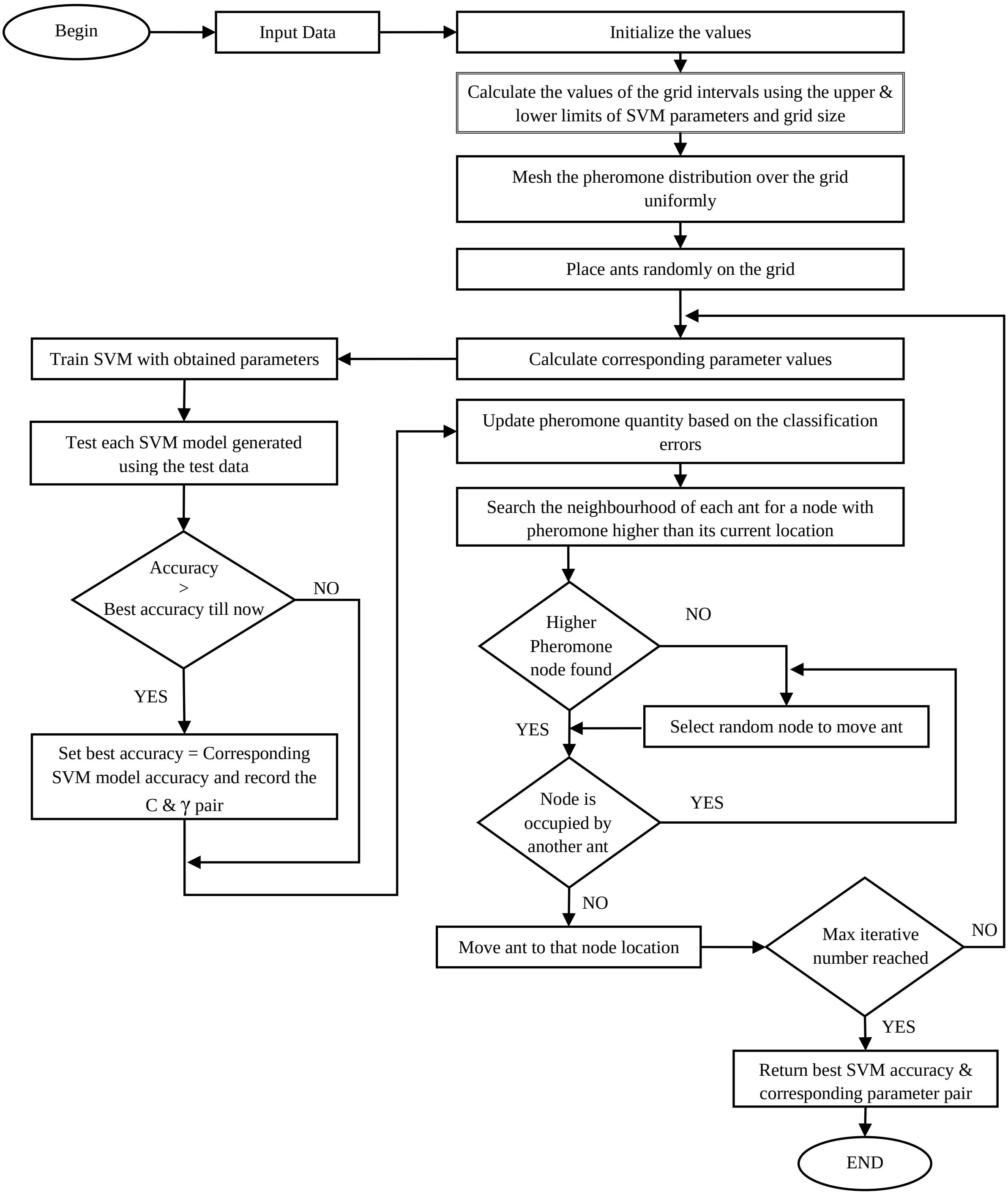

Figure 1. The Process of The Proposed Improved ACO Algorithm for Parameter Optimization of SVM.

IV. EXPERIMENTAL RESULTS

In this section, the performance of the proposed improved ACO algorithm for SVM parameter optimization is assessed. The proposed improved ant optimization algorithm and the technique proposed by Zhang [25] were implemented in computing environment of LIBSVM and C++ Compiler.

Thyroid, Heart, Breast Cancer, Titanic and Diabetes datasets from UCI benchmark and IDA benchmark are used to validate the developed model. The number of attributes for training sample and testing sample are 140 & 75, 170 & 100, 200 & 77, 150 & 2051 and 468 & 300 in Thyroid, Heart, Breast Cancer, Titanic and Diabetes datasets [25] respectively. The number of dimensions in Thyroid, Heart, Breast Cancer, Titanic and Diabetes datasets are 5, 13, 9, 3 and 8 respectively.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Search was performed between 2^{-10} and 2^{10} for parameters C and γ respectively. The key parameters here are number of ants, iterations and grids. The number of ants considered here are 10, 20, 30and 40. The search space for number of iterations and grids ranged from 10 to 100. Initially the evaporation coefficient is predefined at 5 and a timer is set to 0. In the experiment, the timer gets incremented for every iteration of pheromone at grids and if the timer is equal to evaporation coefficient then both are again reset to starting values.

The results obtained are presented in table 1. After experimenting with different number of ants and search space, the proposed algorithm has converged at C=65.368 & γ=2.550 for thyroid dataset, C=0.109 & γ=0.109 for heart dataset, C=0.55 & γ=0.28 for breast cancer dataset, C=6.094 & γ=0.149 for titanic dataset and C=0.541 & γ=0.1 for diabetes dataset. The obtained results are presented below in table 1:

Table 1

Data	Zhang et al. technique				Improved ACO algorithm			
	Optimal C	Optimal γ	Accuracy (%)	Time Taken(s)	Optimal C	Optimal γ	Accuracy (%)	Time Taken(s)
Thyroid	37.38	2.22	97.33	666.20	65.368	2.550	97.999	13.218
Heart	1043.80	2.00	84	519.58	0.109	0.109	81.7	83.407
Breast Cancer	1.98	14.41	74.03	1437.80	0.55	0.28	79.2208	80.528
Titanic	1045.00	54.50	77.43	429.22	6.094	0.149	78.26423	84.915
Diabetes	176.23	151.58	77	19298.00	0.541	0.1	77.3333	107.539

The accuracy in the above table is calculated based on the number of perfectly classified samples. Comparing the results of two methods given in table 1, it can be noticeably observed that the accuracy of the proposed algorithm is high in the case of thyroid, breast cancer, titanic and diabetes dataset, whereas the accuracy in the case of heart dataset is low when compared to the existing technique. These obtained results prove the effectiveness of the improved ACO algorithm.

The best accuracy throughout the experiment was observed with thyroid dataset of 97.999%, where the number of grids are 70 & 80, number of ants are 10 and number of iterations are 10. Different values of the parameters are analysed for all datasets and as illustration, the graphs of titanic are presented below. Firstly, the ranges of C and γ are chosen from 0.1 to 1 and the number of grids and ants are selected to be 10 while iterations are varied from 10, 20, 30 …, 100. Figure 2 demonstrates that as number of iterations increases, the time taken increased. When compared the accuracy with number of iterations, γ & C, it's been observed that accuracy is good in iterations 20, 50, 60, 70, 90 & 100 and C & γ are constant at these positions as shown in figure 3, 4 & 5. Secondly, number of ants and iterations are fixed at 10 and grids are varied from 10 to 100 in equal intervals. Its been perceived in figure 6 that, when number of grids equal to 40, the time taken is less compared to remaining. Figure 7, 8 & 9 shows the accuracy verses number of grids, γ and C respectively, where minimum γ and C are at 90 & 80 grids respectively. Finally, numbers of ants are tested for 10, 20 and 30 and rest of the parameters are fixed at 10. As ants increased, time taken also increased and this is illustrated in figure 10.The figure 11 shows, higher accuracy is observed at 20 and 30 ants and C and γ are identical for those accuracies.

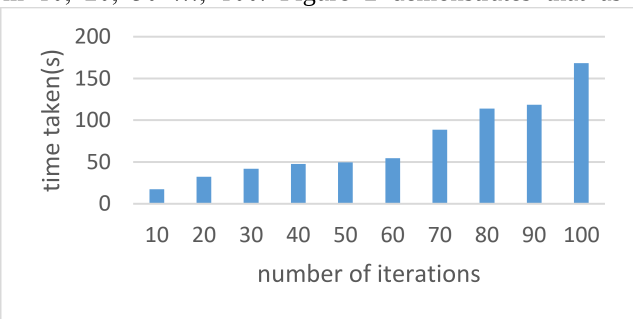

Figure.2. Time Taken by IACO while varying number of number of iterations

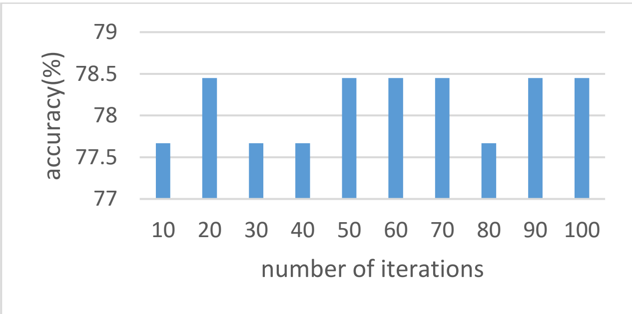

Figure.3. Effect of accuracy of IACO while varying iterations

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

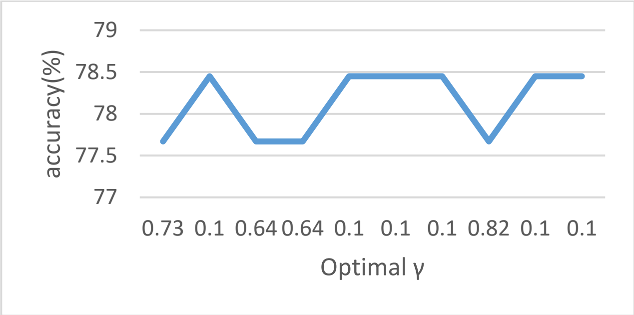

Figure.4. Effect of accuracy of IACO with optimal γ while increasing no. of iterations

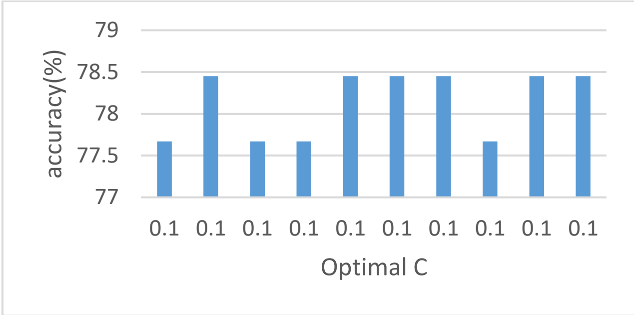

Figure.5. Effect of accuracy of IACO with optimal C while increasing no. of iterations

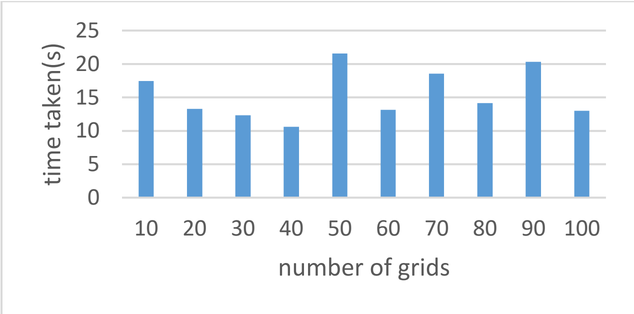

Figure.6. Time Taken by IACO while increasing number of grids.

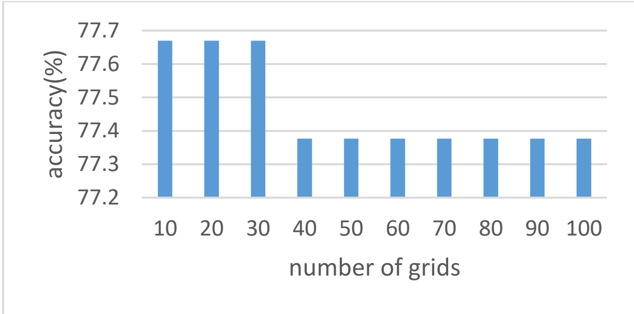

Figure.7. Effect of accuracy of IACO while increasing number of grids

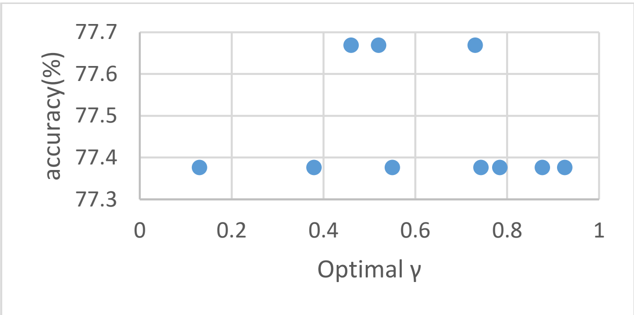

Figure.8. Effect of accuracy of IACO with optimal γ while while increasing no. of grids

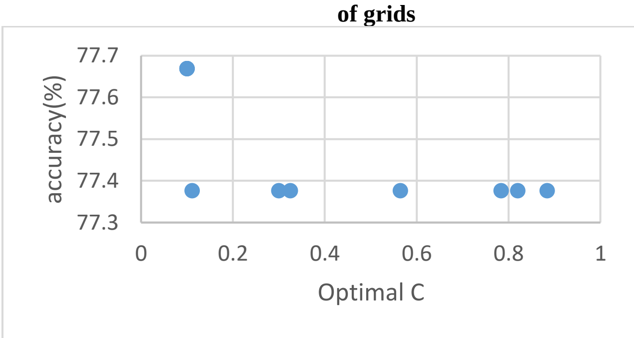

Figure.9. Effect of accuracy of IACO with optimal C increasing no. of grids

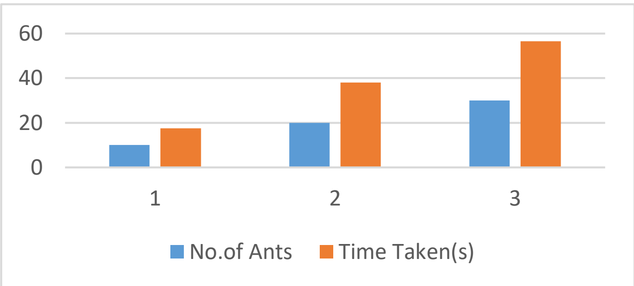

Figure.10. Effect of no. of ants on Time Taken

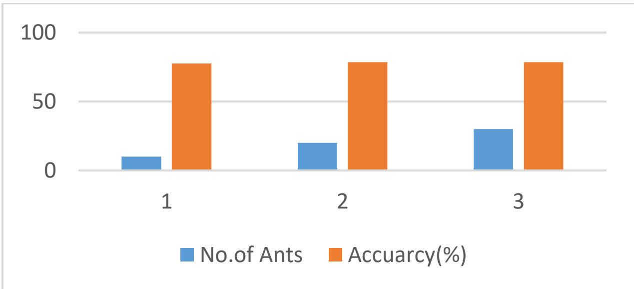

Figure.11. Effect of no. of ants on accuracy

Comparing the results of two methods in table 1 above, it is clear to see that the accuracies of datasets Breast Cancer, Diabetes, Thyroid and Titanic of Improved ACO are more than ACO-SVM algorithm. These results endorse the feasibility and efficiency of the improved ACO algorithm for optimizing the parameters of the SVM.

V. CONCLUSIONS

This paper presents an improved ACO-SVM technique to obtain optimal model parameters. The proposed improved ant colony optimization technique works on overcoming certain disadvantages of the ACO-SVM technique proposed by Zhang et al. like reducing the search space after each iteration and no assurance to the next node when the ant proceeds from one position to the other. The proposed improved model has been tested on standard UCI and IDA benchmark databases. The obtained results demonstrate that proposed method is promising.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

REFERENCES

1. Vapnik VN. The Nature of Statistical Learning Theory. New York, NY: Springer New York; 1995. doi:10.1007/978-1-4757-2440-0.
2. Li X, Zheng A, Zhang X, Li C, Zhang L. Rolling element bearing fault detection using support vector machine with improved ant colony optimization. Measurement 2013;46:2726–34. doi:10.1016/j.measurement.2013.04.081.
3. Samanta B, Al-Balushi KR, Al-Araimi SA. Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection. Eng Appl Artif Intell 2003;16:657–65. doi:10.1016/j.engappai.2003.09.006.
4. Hou S, Li Y. Short-term fault prediction based on support vector machines with parameter optimization by evolution strategy. Expert Syst Appl 2009;36:12383–91. doi:10.1016/j.eswa.2009.04.047.
5. Wu Q, Law R, Wu E, Lin J. A hybrid-forecasting model reducing Gaussian noise based on the Gaussian support vector regression machine and chaotic particle swarm optimization. Inf Sci (Ny) 2013;238:96–110. doi:10.1016/j.ins.2013.02.017.
6. Wu Q. A hybrid-forecasting model based on Gaussian support vector machine and chaotic particle swarm optimization. Expert Syst Appl 2010;37:2388–94. doi:10.1016/j.eswa.2009.07.057.
7. Liu S, Tai H, Ding Q, Li D, Xu L, Wei Y. A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction. Math Comput Model 2013;58:458–65. doi:10.1016/j.mcm.2011.11.021.
8. Wu CH, Tzeng GH, Goo YJ, Fang WC. A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy. Expert Syst Appl 2007;32:397–408. doi:10.1016/j.eswa.2005.12.008.
9. Patil SG, Mandal S, Hegde A V. Genetic algorithm based support vector machine regression in predicting wave transmission of horizontally interlaced multi-layer moored floating pipe breakwater. Adv Eng Softw 2012;45:203–12. doi:10.1016/j.advengsoft.2011.09.026.
10. Chen F, Tang B, Chen R. A novel fault diagnosis model for gearbox based on wavelet support vector machine with immune genetic algorithm. Meas J Int Meas Confed 2013;46:220–32. doi:10.1016/j.measurement.2012.06.009.
11. Lin S-W, Lee Z-J, Chen S-C, Tseng T-Y. Parameter determination of support vector machine and feature selection using simulated annealing approach. Appl Soft Comput 2008;8:1505–12. doi:10.1016/j.asoc.2007.10.012.
12. Huang CL, Wang CJ. A GA-based feature selection and parameters optimizationfor support vector machines. Expert Syst Appl 2006;31:231–40. doi:10.1016/j.eswa.2005.09.024.
13. Aydin I, Karakose M, Akin E. A multi-objective artificial immune algorithm for parameter optimization in support vector machine. Appl Soft Comput 2011;11:120–9. doi:10.1016/j.asoc.2009.11.003.
14. Zhao H, Yin S. Geomechanical parameters identification by particle swarm optimization and support vector machine. Appl Math Model 2009;33:3997–4012. doi:10.1016/j.apm.2009.01.011.
15. Eitrich T, Lang B. Efficient optimization of support vector machine learning parameters for unbalanced datasets. J Comput Appl Math 2006;196:425–36. doi:10.1016/j.cam.2005.09.009.
16. Wang Z, Shao YH, Wu TR. A GA-based model selection for smooth twin parametric-margin support vector machine. Pattern Recognit 2013;46:2267–77. doi:10.1016/j.patcog.2013.01.023.
17. Zhao M, Fu C, Ji L, Tang K, Zhou M. Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes. Expert Syst Appl 2011;38:5197–204. doi:10.1016/j.eswa.2010.10.041.
18. Reif M, Shafait F, Dengel A. Meta-learning for evolutionary parameter optimization of classifiers. Mach Learn 2012;87:357–80. doi:10.1007/s10994-012-5286-7.
19. Zhang Q, Shan G, Duan X, Zhang Z. Parameters optimization of Support Vector Machine based on Simulated Annealing and Genetic Algorithm. 2009 IEEE Int Conf Robot Biomimetics 2009:1302–6. doi:10.1109/ROBIO.2009.5420717.
20. Adankon MM, Cheriet M. Genetic algorithm-based training for semi-supervised SVM. Neural Comput Appl 2010;19:1197–206. doi:10.1007/s00521-010-0358-8.
21. Boardman M, Trappenberg T. A Heuristic for Free Parameter Optimization with Support Vector Machines. 2006 IEEE Int Jt Conf Neural Netw Proc 2006:1337–44. doi:10.1109/IJCNN.2006.246739.
22. Yu Q, Zhang B, Wang J. Parameter Optimization of ϵ-Support Vector Machine by Genetic Algorithm. 2009 Fifth Int Conf Nat Comput 2009:540–2. doi:10.1109/ICNC.2009.628.
23. Huaitie X, Guoyu F, Zhiyong S, Jianjun C. Hybrid Optimization Method for Parameter Selection of Support Vector Machine. 2010 IEEE Int Conf Intell Comput Intell Syst 2010;1:613–6. doi:10.1109/ICICISYS.2010.5658455.
24. Lin S-W, Ying K-C, Chen S-C, Lee Z-J. Particle swarm optimization for parameter determination and feature selection of support vector machines. Expert Syst Appl 2008;35:1817–24. doi:10.1016/j.eswa.2007.08.088.
25. Zhang XL, Chen XF, He ZJ. An ACO-based algorithm for parameter optimization of support vector machines. Expert Syst Appl 2010;37:6618–28. doi:10.1016/j.eswa.2010.03.067.
26. Alwan HB, Ku-Mahamud KR. Optimizing Support Vector Machine Parameters Using Continuous Ant Colony Optimization. Proc 7th Int Conf Comput Converg Technol 2012:164–9.
27. Zhang XL, Chen XF, Zhang ZS, He ZJ. A Grid-based ACO Algorithm for Parameters Optimization in Support Vector Machines. IEEE Int Conf Granul Comput 2008 GrC 2008 2008:805–8. doi:10.1109/GRC.2008.4664645.
28. Vapnik VN. An overview of statistical learning theory. IEEE Trans Neural Netw 1999;10:988–99. doi:10.1109/72.788640.
29. Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. 1st ed. Cambridge: Cambridge University Press; 2000.
30. Jiawei H, Kamber M. Data mining: concepts and techniques. 2006. doi:10.1007/978-3-642-19721-5.
31. Duan K, Keerthi S, Poo A. Evaluation of simple performance measures for tuning SVM hyper parameters. Neurocomputing 2003;51:41–59.
32. Imbault F, Lebart K. A Stochastic Optimization Approach for Parameter Tuning of Support Vector Machines. Proc 17th Int Conf Pattern Recognition, 2004ICPR 2004 2004;4:597–600. doi:10.1109/ICPR.2004.1333843.
33. Chapelle O, Vapnik V, Bousquet O, Mukherjee S. Choosing Multiple Parameters for Support Vector Machines. Mach Learn 2002;46:131–59. doi:10.1023/A:1012450327387.

R. Srujana: received her B.Tech & M.Tech (Computer Science & Information Technology) dual degree from Andhra University. She is currently a Research Scholar in the Department of Computer Science & Engineering at GITAM University. She is working for her PhD in the area of Model Selection. Her current research interest includes Data Mining, Computational Intelligence and Machine Learning.

Dr. Y. Radhika: received her B.E (Civil Environmental Engineering), M.Tech (Computer Science and Technology) and PhD degrees from Andhra University. She is currently a Professor in the Department of Computer Science & Engineering at GITAM University. She has been publishing in national and international journals. Her current research interest includes Data Mining, Neural Networks, etc.