A Mutated Salp Swarm Algorithm for Optimization of Support Vector Machine Parameters

R.R.Rajalaxmi
Department of Computer Science and Engineering,
Kongu Engineering College,
Perundurai -638060,
rrrkec.69@gmail.com

E.Vidhya
Department of Computer Science and Engineering,
Kongu Engineering College,
Perundurai -638060,
vidhya.eswaramurthy@gmail.com

***Abstract*-Support Vector Machine (SVM) is typically a supervised learning algorithm that carefully examines input and identifies distinct patterns. The function of SVM classifier relies on adjusting or controlling of kernel and penalty parameter values. Nature Inspired Algorithm helps to solve the natural problems and has been attracting considerable attention due to their better performance. Salp Swarm Algorithm (SSA) is a Nature Inspired Algorithm (NIA) which is used to control the finest SVM parameters value. To improve exploration capability of SSA, mutation method is developed to find the optimal value for kernel parameter and penalty parameter. The preliminary result indicates Mutated SSA with SVM increases classification accuracy than simple SSA with SVM.**

Keywords-Support Vector Machine, Nature Inspired Algorithm, Salp Swarm Algorithm, Mutation, Kernel parameter, Penalty parameter.

I. INTRODUCTION

Data mining is a technique for recognizing hidden pattern to evaluate large amounts of data from databases or data warehouse [1]. Classification is an effective data mining approach, which helps to predict group membership for data object. The intent of classification is to accurately predict the target class for each case in the data. SVM is a supervised learning mechanism applied for classification. SVM can be applied for numerical prediction, data interpretation, and recognizes the pattern [1]. The goal of SVM is used wisely to construct one or more hyperplane to partition various distinct classes. It facilitates to recognise the optimal hyperplane that enlarges the margin within two different classes. Support vectors represent the points that reside on this margin. A support vector is the point from each class that are nearest to the maximum margin hyperplane. Each class have at least one support vector.

SVM can operate even in infinite dimensions. If the instance is linearly separable, then SVM produces high accuracy rate. SVM performs low, when the instance is non-linearly separable. On correlating the sample from the input element into a high-dimensional element, then kernel function aids to solve the non-linearly separable problem. A main challenge of SVM is adjusting their parameters and selection of suitable kernel function. SVM prevent from over-fitting and it can be work with relatively larger number of feature without requiring too much computation.The performance of SVM can be accomplished by utilizing kernel function and penalty parameter (C) [2].

A. Penalty Parameter

Penalty parameter (C) is a SVM parameter to resolve the constrained optimization problem [3, 4]. In SVM, the penalty parameter C manipulates the error on the training data and maximizing the margin. It is also known as Regularization parameter. When C has small value, then it will increase the training error and get larger margin hyperplane. When C has high value, it will decrease the training error and get smaller margin hyperplane. The Penalty parameter is used to avoid overfitting.

B. Kernel Function

This function plays a major part in SVM. The kernel functions are depleted once the samples are non-linearly separable. The facts of kernel function are used to allow operations to be depicted within the input element rather than the potentially high dimensional element. Instead of computing the mapping explicitly, the samples will gain meaningful linear separable data. “Kernel Trick” is a process of converting non-linearly separable instance in an input space into a linearly separable instance in the element [4].

Some challenges in kernel function are determining optimum value for a given problem with minimum complexity of kernel function for accessing the element with the help of` relevant input samples. There are different types of kernel functions namely polynomial kernel, linear kernel, sigmoid kernel, and Radial Basis Function (RBF).

1) *Linear Kernel:* This kernel is the most straightforward kernel function given by the dot product. It is often related to non-kernel counterparts.

$$K(X_i, X_j) = X_i \bullet X_j \tag{1}$$

2) *Polynomial Kernel:* This is the commonly used kernel function in SVM [4] and also a universal method for non-linear modeling. Polynomial kernel contains all polynomials terms up to degree d. These kernels are well suited for normalized training data. It is also known as non-stationary kernel.

$$K(X_i, X_j) = (X_i \bullet X_j + 1)^d \tag{2}$$

978-1-5386-9533-3/19/$31.00 ©2019 IEEE

3) *Radial Basis Function Kernel:* RBF kernel is a well-known kernel function used frequently in SVM. This kernel non-linearly projects the instances over a high dimensional space separating from the linear kernel. It is also known as Gaussian Kernel. It has limited hyper parameters than the polynomial kernel with less numerical difficulties.

The RBF kernel enables to make a very complex decision boundary on a high dimensional feature mapping with efficient computation due to the kernel representation. σ is a hyper parameter and also referred to as the Kernel Bandwidth [4]. When σ is small, it will have a very fine-tuned boundary and larger estimation error. When σ is large, then the function has a very smoother boundary with large approximation error.

$$K(X_i, X_j) = e^{\left(-\frac{\|X_i - X_j\|^2}{2\sigma^2}\right)} \tag{3}$$

4) *Sigmoid Kernel:* The Sigmoid Kernel referred as Hyperbolic Tangent Kernel originates from neural network environment which utilizes the bipolar sigmoid capacity as an activation function for artificial neurons. Thus, it resembles similar to two-layer neural network known as a Multilayer Perceptron (MLP) kernel without hidden layer.

$$K(X_i, X_j) = \tanh(\alpha X_i X_j + C) \tag{4}$$

Alpha (α) and intercept constant C are the most commonly utilized adjustable parameters. A typical estimation of alpha is 1/N, where N denotes the dimension of data [4].

II. Literature Review

Suitable SVM parameter selection helps to enhance accuracy rate with minimum classification error. The optimal value of the SVM parameters can be obtained using Salp Swarm Algorithm with improved classification accuracy. New techniques for pattern classification that effectively set the parameters of a kernel for SVM during the training phase leads to improved accuracy. Genetic Algorithm (GA) is one method to improve the parameters to pick the most effective subset of features while not diminishing the SVM classification accuracy [5]. Another algorithm namely Particle Swarm Optimization (PSO) with SVM approach chooses the Kernel parameters for SVM together with feature selection [6]. The algorithm finds the value of parameters with relevant features and without diminishing the classification accuracy. This methodology yields better classification accuracy rate than other methods similar to GA with SVM.

In another research, Hybrid Genetic Algorithm with Support Vector Regression (HGA-SVR) method [7] is developed to solve complex forecasting problem in an effective way and also increase the prediction rate of Support Vector Regression (SVR). This method identified a suitable kernel function and also optimized the parameter value of the selected kernel with reduced prediction error in electricity load forecasting. An Ant Colony Optimization with Support Vector Machine (ACO-SVM) [8] is applied over few benchmark dataset to assess the feasibility and efficiency. This algorithm identifies the best parameter to establish an effective SVM and also achieves adequate accuracy level. In another research, Genetic Algorithm [9] with feature selection for SVM classification is adapted to improve accuracy with lesser processing time and also select relevant feature subsets. The algorithm also determines the best parameters of SVM. In standard SVM methods, penalty parameter C and kernel function are typically based on experience. Finding an optimal parameter value over different data with high accuracy is difficult. To address these challenges, the GA-SVM forecast model [10] is presented to enhance the parameter selection of SVM.

A new novel hybrid algorithm namely Bat Algorithm with Harmony Search (HS/BA) is developed for solving numerical problems. The proposed algorithm converges fast and also provides feasible solution for real-life problem [11]. In another work, hybridized meta-learning with multi-objective PSO is developed to finalize the SVM parameters [12]. This algorithm employed ML procedure to propose Pareto front of SVM and multi-objective optimization algorithm with a higher quality of classification accuracy compared to a traditional multi-objective algorithm. Global optimization [13] method enables binary SVM with less training and testing and also identified the best binary tree structure. The proposed method enhances SVM accuracy for multiclass problems of medium and huge size within time. A new hybrid approach has been specifically developed for solving many real-world problems in different areas.

Nonparallel Support Vector Machine (NPSVM) is more flexible and more capable than Twin-SVMs (TWSVMs). One of the primary challenges in executing NPSVM is tuning particular parameters throughout the training procedure [14]. PSO is used for parameter specification in NPSVM with less run time and good in comparison to TWSVM and NPSVM. Hypergraph based Genetic Algorithm (HG-GA) is another approach to set the parameters and feature selection in SVM. HG-GA effectively searched the best solution to avoid the local minima issue for excellent classifier accuracy [15]. Genetic Algorithm with Support Vector Machine is designed to deal with FOREX trading. SVM is used to identify the type of market with Genetic Algorithm to make perfect rules in each market [16]. Based on the prediction, dynamic optimized algorithm is able to achieve trading rules and to make assets with good leverage quality. The preliminary results of this algorithm reveal that the 85% of Return on Investment (ROI). A new methodology called Bat Algorithm with Support Vector Machine (BA-SVM) is utilized to compute the optimal values of SVM parameters and minimum classification error. This algorithm prevents the local optima problem and it is tested with UCI repository datasets. This method yields minimum classification error than PSO and GA.

III. Nature Inspired Algorithm

Nature inspired computing is a technique inspired by processes, observed from nature and these techniques mainly lead to development of algorithms called Nature Inspired Algorithms (NIAs). The purpose of developing the

NIAs is to address complex real-world optimization problems, to optimize engineering problems, multi objective functions and also deal with NP-hard problems for large number of variables [17]. NIAs solve the problems in an efficient and effective manner and also provide an optimum solution.

Many nature inspired algorithms are used for optimizing the search operation, prediction, classification, clustering and neural networks. Nature has driven several researchers to develop different algorithms. Nature inspired algorithms are categorized into four types namely, Swarm Intelligence (SI) approach, Bio-inspired (BI) approach, Evolutionary approach, Physics and Chemistry approach, and other approaches.

SI is an artificial intelligence based method formed from the cumulative behavior of natural, decentralized, and self-organized systems. The SI algorithm is stimulated by the swarm behavior of distinct species like birds, fishes, ants, bees, fireflies, cats, bats, etc., Many algorithms emerged using Swarm Intelligence system: Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Bat Algorithm (BA), Cuckoo Search Algorithm (CSA), Firefly Algorithm (FA), Salp Swarm Algorithm (SSA) etc.,

IV. SALP SWARM ALGORITHM

SSA is the latest SI based algorithm [4, 18]. It adopts mimicking behaviour of salp in sea and their social interaction. Salps are usually floated together as salp chain to give finer movement by quick reconciliation and foraging in oceans. It has good exploration and exploitation capabilities to find an optimal solution.

Salp Swarm Algorithm is a population based method. Salps come under the Salpidae family and it may look like a Spineless jellyfish [19]. Salps tissues are wildly identical to jelly fish and move towards in searching of food. Salp has transparent barrel-shaped body. They are living in deep sea or oceans and grow rapidly. Salps move due to water forces to search their food in ocean to form swarms called as salp chains. The size of an individual sea salp is approximately 10cm long and length of a salp chain can be upwards of 4.5 meters long. SSA is developed to resolve the optimization issue in an effective manner and also help to get struck in local minima.

Salp chain is partitioned into two types such as leader and follower [18]. Salp leader is at frontend of a chain and remaining salps are the follower. The leader monitors the entire swarm and the follower follows leader instruction. Salps position is represented as an n-dimensional search space where n denotes the number of attribute of a problem. The salps position are maintained in two-dimensional matrix named X. The purpose of salp swarm is to determine food source (F) in solution space. The leader's position is modified as per equation:

$$X_j^1 = \begin{cases} F_j + C_1((ub_j - lb_j)C_2 + lb_j & C_3 \geq 0 \\ F_j - C_1((ub_j - lb_j)C_2 + lb_j & C_3 < 0 \end{cases} \quad (5)$$

where,

x_j^1 - Leader of salp position of j^{th} dimension.
F_j - Food source position of j^{th} dimension.
ub_j - Upper bound of j^{th} dimension.
lb_j - Lower bound of j^{th} dimension.

C1, C2, and C3 parameters are generated using random numbers. Parameters C2 and C3 lie in the interval [0,1]. C3 is responsible for next position of current leader salp towards $+\infty$ or $-\infty$. The most effective parameter in SSA is C1 which balances between the exploration and exploitation. C1 is decreased over iteration and can be calculated as per equation (6).

$$C_1 = 2e^{-\left(\frac{4l}{L}\right)^2} \quad (6)$$

where,

l - Present iteration.
L - Maximal number of iteration.

The follower position is updated as per Newton's law using equation (7).

$$X_j^i = \frac{1}{2}at^2 + v_0 t \quad i \geq 2 \quad (7)$$

where,

x_j^i - i^{th} follower salp position in j^{th} dimension.
t - Time
v_0 - Initial speed

a - $\dfrac{v_{final}}{v_0}$

v - $\dfrac{x - x_0}{t}$

In optimization, time represents iteration and, v_0=0. Hence equation (7) can be reformulated as

$$X_j^i = \frac{1}{2}(X_j^i + X_j^{i-1}) \quad i \geq 2 \quad (8)$$

SSA-SVM algorithm cannot obtain better result due to poor exploration of the solution. In order to improve the exploration for the given solution, we proposed a mutated SSA-SVM to get better accuracy. Mutation enables to diversify the solution from one generation to another generation in the population. It helps in search to converge faster by getting different solution. It is possible to avoid premature convergence on a local maximum or minimum. One of the mutation methods is Gaussian mutation as per equation (9).

$$f_{Gaussian(0,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (9)$$

Algorithm: Mutated Salp Swarm Optimization

Input
Dataset;
Maximum number of iteration;
Search Agents number;
Kernel Parameter (σ);
Penalty Parameter (C);
Output
Best accuracy
begin
Initialize salps x_i(i=1, 2, ..., n)with respect to ub and lb
Identify salps with minimum fitness and mutate them as per equation(9)
while (fitness(F)<threshold)
Compute the fitness of each salp as per equation (10)
F=the optimal search agent
Modify C_1 as per Equation (6)
    For each salp (x_i)
        if(i==1)
            Modify leader salp's position as per Equation (5)
        else
            Modify follower salp's position as per Equation (8)
        end
    end
Adjust the salps using the ub and lb
end
return F

Fig. 1. Mutated SSA pseudo code

Fig. 1 depicts the Mutated SSA pseudo code. It initialises multiple salps with random position for optimization process and also performs mutation for salps with minimum fitness. It computes the fitness for each position of salp, determines the optimal fitness and assigns the finest salp's position to F as the food source. C1 is updated using the Equation (6). The leader salp's position is updated for each dimension as per Equation (5).

V. RESULT AND DISCUSSION

The proposed method is tested on datasets available in UCI machine learning repository. Table I depicts the dataset details to test the effectiveness of the algorithm. The performance metric used for testing the selected parameter in SVM is accuracy. Accuracy is the traditional way to measure the performance of the system and it is one of the assessment metric of classification method. It denotes the ratio of accurately classified samples to the total number of samples calculated using Equation (10).

$$\text{Accuracy} = \frac{\text{No. of correctly classified Samples}}{\text{No. of Samples}} \qquad (10)$$

TABLE I. DATASET DETAILS

Dataset	No. of Dimensions	No. of Samples	Classes
Ionosphere	34	351	2
Liver-disorder	6	345	2
Sonar	60	208	2

In order to analyse the results, Mutated SSA is compared with SSA-SVM. The proposed algorithm optimizes the SVM parameters with improved classification accuracy. The parameter values while implementing the Mutated SSA algorithm are shown in Table II.

TABLE II. PARAMETERS FOR IMPLEMENTATION

S.No	Parameter	Value
1	Number of search agents	40
2	Number of iteration	500
3	Dimension	2
4	Number of repetitions of runs	5
5	Penalty parameter (C)	0.01 to 500
6	Kernel parameters (σ)	0.01 to 100

Table III describes the optimal value of kernel parameter (σ) and penalty parameter (C) and their accuracy for Ionosphere dataset.

TABLE III. PARAMETER VALUE AND COMPARISON OF ACCURACY FOR IONOSPHERE DATASET

Ionosphere			
	SVM	***SSA-SVM***	***Mutated SSA-SVM***
Penalty parameter (C)	0.2	88	118
Kernel parameter (σ)	100	100	86
Accuracy (%)	64.67	87.46	91.17

Table IV shows the optimal value of kernel parameter (σ) and penalty parameter (C) and their accuracy for Sonar dataset.

TABLE IV. PARAMETER VALUE AND COMPARISON OF ACCURACY FOR SONAR DATASET

Sonar			
	SVM	***SSA-SVM***	***Mutated SSA-SVM***
Penalty parameter (C)	0.2	64	60
Kernel parameter (σ)	100	100	70
Accuracy (%)	53.37	75.48	82.69

Table V represents the best value of kernel parameter (σ) and penalty parameter (C) and their accuracy for Liver-Disorder dataset.

TABLE V. PARAMETER VALUE AND COMPARISON OF ACCURACY FOR LIVER-DISORDER DATASET

Liver-Disorder			
	SVM	***SSA-SVM***	***Mutated SSA-SVM***
Penalty parameter (C)	0.2	0.01	86
Kernel parameter (σ)	100	85	15

Accuracy (%)	59.52	59.71	76.23

Fig. 2 shows the accuracy obtained using Support Vector Machine, SSA-SVM and Mutated SSA-SVM. From the results, Mutated SSA-SVM has better accuracy when compared to SVM and SSA-SVM.

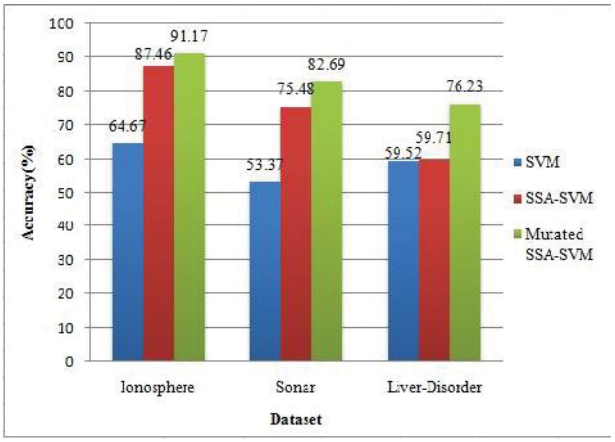

Fig. 2. Accuracy of Mutated SSA-SVM

VI. CONCLUSION AND FUTURE WORK

The major factors to decide the SVM performance are finding the optimal value of kernel (σ) and penalty parameters (C). Nature inspired algorithm is one of the optimization technique which can be provide some efficient ways to solve nature problems. Salp Swarm Algorithm is used for identifying best values of the SVM parameters. Mutated SSA with SVM is used to control the parameter value which gives the effective solution by probing and adopting. The result shows the Mutated SSA with SVM achieved better classification accuracy and avoid local optima problem. In future, the hybridization of SSA or any nature inspired algorithm can be used to optimize the parameter value and try to achieve good classification accuracy.

REFERENCES

[1] J. Han, M. Kamber, J. Pei, " Data Mining: Concepts and Techniques", Third Edition, Morgan Kaufman publishers, 2012.

[2] S. Li, H. Fang, X. Liu, "Parameter optimization of support vector regression based on sine cosine algorithm" Expert systems with Applications, vol. 91, pp. 63-77, August 2017.

[3] A. A. Abusnaina, R. Jarrar, S. Ahmad, M. Mafarja, "Training neural networks using salp swarm algorithm for pattern classification" 2018.

[4] R. Amami, D. B. Ayed, N. Ellouze, "Practical selection of SVM supervised Parameters with different feature representations for vowel recognition", JDCTA, vol. 7, May 2013.

[5] V. Arulkumar. "An Intelligent Technique for Uniquely Recognising Face and Finger Image Using Learning Vector Quantisation (LVQ)-based Template Key Generation," International Journal of Biomedical Engineering and Technology 26, no. 3/4 (February 2, 2018): 237-49.

[6] C-L. Huang, C-J. Wang, "A GA-based feature selection and parameters optimization for support vector machines" Expert Systems with Applications, vol. 31, pp. 231-240, 2006.

[7] C.V. Arulkumar, G. Selvayinayagam and J. Vasuki, "Enhancement in face recognition using PFS using Matlab," International Journal of Computer Science & Management Research, vol. 1(1), pp. 282-288, 2012

[8] S-W. Lin, K-C. Ying, S-C. Chen, Z-J. Lee, "Particle swarm optimization for parameter determination and feature selection of support vector machines" Expert Systems with Applications, vol. 35, pp. 1817-1824, 2008.

[9] C-H. Wu, G-H. Tzeng, R-H. Lin, "A Novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression" Expert Systems with Applications, vol. 36, pp. 4725-4735, 2009.

[10] X. Zhang, X. Chen, Z. He, "An ACO-based algorithm for parameter optimization of support vector machines" Expert Systems with Applications, vol. 37, pp. 6618-6628, 2010.

[11] C. V. Arulkumar, P. Vivekanandan, "Multi-feature based automatic face identification on kernel eigen spaces (KES) underunstable lighting conditions", Advanced Computing and Communication Systems 2015 International Conference on. IEEE, 2015.

[12] M. Zhao, C. Fu, L. Ji, K. Tang, M. Zhou, "Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes" Expert Systems with Applications, vol. 38, pp. 5197-5204, 2011.

[13] X. Guo, D. Li, A. Zhang, "Improved Support Vector Machine Oil Price Forecast Model Based on Genetic Algorithm Optimization Parameters" AASRI Procedia 1, pp. 525-530, 2012.

[14] G. Wang, L. Guo, "A Novel Hybrid Bat Algorithm with Harmony Search for Global Numerical Optimization" Journal of Applied Mathematics, vol. 696491, December 2012.

[15] P. B. C. Miranda, R. B. C. Prudencio, A. P. L. F. Carvalho, "A hybrid meta-learning architecture for multi-objective optimization of SVM parameters" Neurocomputing, vol. 143, pp. 27-43, June 2014.

[16] Y. Lee, J. Lee, "Binary tree optimization using genetic algorithm for multiclass support vector machine" Expert Systems with Applications, vol. 42, pp. 3843-3851, January 2015.

[17] S. M. H. Bamakan, H. Wang, A. Z. Ravasan, "Parameters optimization for Nonparallel Support Vector Machine by Particle Swarm Optimization" Procedia Computer Science, vol. 91, pp. 482-491, 2016.

[18] M. R. Gauthama Raman, N. Somu, K. Kirthivasan, R. Liscano, S. V. S. Sriram, "An efficient intrusion detection system based on hypergraph – Genetic algorithm for parameter optimization and feature selection in support vector machine" knowledge-Based Systems, vol. 134, pp. 1- 12, July 2017.

[19] Anandakumar, "Energy Efficient Network Selection Using 802.16g Based Gsm Technology," Journal of Computer Science, vol. 10, no. 5, pp. 745–754, May 2014.

[20] B. J. Almeida, R. F. Neves, N. Horta, "Combining Support Vector Machine with Genetic algorithms to optimize investments in Forex markets with high leverage" Applied Soft Computing, vol. 64, pp. 596-613, January 2018.

[21] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, S. M. Mirjalili, "Salp swarm algorithm: A bio-inspired optimizer for engineering design problems" advances in Engineering Software, vol. 114, pp. 163-191, July 2017.

[22] www.alimirjalili.com/SSA.html

[23] H. Faris, M. M. Mafarja, A. A. Heidari, I. Aljarah, A. M. Al-Zoubi, S. Mirjalili, H. Fujita, "An efficient binary salp swarm algorithm with crossover scheme for feature selection problems" Knowledge-Based Systems, vol. 154, pp. 43-67, May 2018.

[24] C.V. Arulkumar et al., "Secure Communication in Unstructured P2P Networks based on Reputation Management and Self Certification", International Journal of Computer Applications, vol. 15, pp. 1-3, 2012.

[25] T-J. Hsieh, H-F. Hsiao, W-C. Yeh, "Mining financial distress trend data using penalty guided support vector machines based on hybrid of particle swarm optimization and artificial bee colony algorithm" Neurocomputing, vol. 82, pp. 196-206, December 2011.

[26] X. Li, A. Zheng, X. Zhang, C. Li, Li. Zhang, "Rolling element bearing fault detection using support vector machine with improved ant colony optimization" Measurement, vol. 46, pp. 2726-2734, May 2013.