

A Parameter Selection Method of Support Vector Machine Based on Improved FOA

Shi Huishu^{1, a *}, Zhu Yi^{1, b}, San Ye^{1, c} and Wang Lei^{2, d}

¹Control and Simulation Center Harbin Institute of Technology, No.92 West Straight Street, Harbin, Heilongjiang, China

²Aviation Repair Factory 61255 Unit of People Liberation Army Houma, Shanxi, China

^ashihuishu_10@126.com, ^bzhuyi@hit.edu.cn, ^csanye@hit.edu.cn, ^dwanglei_sw@126.com

Keywords: Parameter selection; Fruit Fly Optimization algorithm; Support vector machine; Uniform design

Abstract. Parameter selection is an important issue for support vector machine. To solve that problem, an improved fruit fly optimization algorithm is proposed in this paper, and used for parameter selection of support vector machine. The initial position and the iteration step size are improved in the new algorithm, and it has stronger search capability. Effectiveness and generality of the method are proved by experiment on standard data sets.

Introduction

Support Vector Machine (SVM), which is based on structural risk minimization and VC dimension, has significant advantage in classification. As one of the most popular method in machine learning, SVM is widely used in handwriting recognition, face recognition, text categorization, signal processing and fault diagnosis.

The parameter has an important effect on the performance of SVM, which make parameter selection a hot issue of SVM research. There are 2 main methods on parameter selection: grid search [1] and intelligent algorithms [2, 3]. Grid search has advantages of reliability and easy to use, but it is essentially an exhaustive method and requires a lot of computation. Especially in the case of parameter multifaceted, the parameter search would be very difficult. Intelligent algorithms have been concerned for nonlinear optimization capability, but the parameters of themselves have a certain effect on SVM parameter search. Furthermore, the shortcomings of intelligent methods like local extremum and computation complexity, limit the application on SVM parameter selection.

Fruit fly optimization algorithm (FOA) is a new optimization algorithm, which has fewer parameters and higher computational efficiency than the other traditional intelligent algorithm. An improved FOA is proposed in this paper to solve the SVM parameter selection problem.

Fruit Fly Optimization Algorithm

FOA is a new swarm intelligent algorithm proposed by W. T. Pan [4]. The algorithm considers the food of fruit fly as optimal solution and search the optimal solution by imitating the foraging of fruit fly. The steps of FOA are as follows:

- 1) Set the size of the fruit fly population and the initial position (X_0, Y_0) .
- 2) Assign the search direction and the search distance of fruit fly randomly.

$$X_i = X_0 + \text{Random}, Y_i = Y_0 + \text{Random} . \quad (1)$$
- 3) Calculate the distance D between fruit fly and the origin, the odor concentration discriminant value S of each fruit fly, and the position odor concentration Smell_i .

$$D_i = \sqrt{X_i^2 + Y_i^2}, S_i = 1/D_i, \text{Smell}_i = \text{fitness}(S_i) . \quad (2)$$
- 4) The position of fruit fly which has the highest concentration in the population, is considered as the optimal position, the rest fruit flies flight to that position.

5) Repeat the steps 2-4. If $Smell_i$ of new generation is better than previous generation, retain the new optimal position. Otherwise, retain the optimal position of the previous generation.

Initial position, population size and flight distance mainly affect the convergence speed of FOA. Usually, a good initial position and large population size can make the algorithm converges fast. Increase flight distance can improve search speed.

For SVM parameter selection, the parameter to be searched was represented by S and the classification accuracy of SVM with the parameters was represented by $Smell_i$.

SVM Parameter Selection Method Based on Improved FOA

SVM parameter selection is a multi-extremum problem. Search SVM parameter by original FOA is easy to fall into local extremum and difficult to obtain the optimal parameter.

On the other hand, the flight distance of fruit fly, also known as iterative step size, has a significant effect on optimization capability of the algorithm. A large step size is conducive to jump out of local extremum, but reduce local search capability; a small step size may fall into local extremum, and reduce search efficiency. For that reason, a fixed iterative step size is difficult to adapt the entire optimization process. In addition, SVM classification performance has different sensitivity to different parameters and the same step size is obviously not suitable for all parameters.

To solve the above issues, an improved FOA with fixed initial position and variable step size is proposed in this paper.

Fixed Initial Position Initial position of fruit fly population is random, and it is replaced by a given position in this paper to improve the search capability and efficiency of the algorithm.

Uniform design (UD) is a direct optimization method, which follows the principle of uniform dispersion. The key of the method is to make most of uniform points to test, and the aim is to get the good result using less test points. It is an effective way to reduce the amount of experiments. For that reason, UD is suitable for test analysis of complex systems with multi-level and multi-factor. UD has higher efficiency than grid search.

In this paper, UD is used to search parameter scope, and then the parameters which make SVM performance best from the experimental results are selected to be the initial position of the fruit fly population. By giving an initial position with good performance, search efficiency of the algorithm can be improved and the risk of falling into local extremum can be reduced.

Variable Step Size To reduce the effect of parameter sensitivity, fixed iterative step size is replaced by variable step size. Make iterative step size correlated with the parameters of SVM, and large step size is used at first to improve the search speed, small step size is used to enhance the local search capability of the algorithm.

As can be seen in Eq. 1, the search direction and distance is random, let the interval of iterative step size is $[-r, r]$, express the random value as

$$Random = 2r * rand() - r. \quad (3)$$

Obviously, fruit fly individual flight direction and distance control by the parameter r .

The iterative step size of the fruit fly should correlate with the SVM parameter, which is express

as the odor concentration discriminant value S . From Eq. 2, there is $-\frac{1}{S_i} \leq X_i \leq \frac{1}{S_i}, -\frac{1}{S_i} \leq Y_i \leq \frac{1}{S_i}$.

Simply, let

$$r = k \cdot \frac{1}{S_i}, k \in (0.05, 0.5), \quad (4)$$

and the range of the random value is $\left[-\frac{k}{S_i}, \frac{k}{S_i}\right]$.

According to SVM classification accuracy and iteration times, adjust k to $\alpha \cdot k$ in the search process. The basic principles are as follows:

- 1) If the classification accuracy of SVM unchanged after n times, let $\alpha > 1$ to jump out of local extremum.
- 2) If the classification accuracy of SVM increased significantly, or in the late of search process, let $0.5 < \alpha < 1$ to improve the local search capability.

Experiment and analysis

Experimental Design and Data Sets The core of SVM is kernel function. To verify the universality and effectiveness of the improved FOA, Gaussian kernel, sigmoid kernel, polynomial kernel and combination kernel function based on Gaussian kernel and polynomial kernel are chosen to constitute SVM. The SVMs were used to classify 5 standard data sets from UCI database in this section. The data sets were iris; wine; heart; vehicle and waveform. All the data are normalized before the experiment.

It should be note that the combination kernel function is

$$\kappa(x, z) = \alpha \cdot \exp(\gamma_{rbf} \|x - z\|^2) + (1 - \alpha) \cdot (\gamma_{poly}(x, z) + c)^d \quad (5)$$

where $\alpha \in (0.2, 0.9)$ is a proportionality coefficient.

Experiment and Analysis For all SVMs but the combination-SVM, the value range was divided into 15 levels according to $U_{15}^*(15^7)$ table. The value range of combination kernel SVM was divided into 16 levels according to $U_{16}^*(16^{12})$ table. UD tables can be found in reference [5]. The performance of the SVM parameters was evaluated by 10-fold cross validation.

The initial position of the fruit fly population is obtained by UD, let population size is 20 and the number of iterations is 50. For the sensitive parameters like γ , let $k = 0.1$. For the insensitive parameters like C , let $k = 0.3$. In this paper, let $\alpha = 1.2$ when the number of iteration is 5, and let $\alpha = 0.75$ after iteration times is 40.

The experimental results are shown in Table 1 and Table 2.

Table 1 Classification accuracy of SVMs on different data sets (%)

	Iris	Wine	Heart	Vehicle	Waveform
Gauss-SVM	97.33	98.88	84.81	86.64	87.20
Sigmoid-SVM	97.33	98.88	84.44	81.56	87.08
Polynomial-SVM	97.33	98.88	84.81	86.41	87.12
Combination-SVM	97.33	98.88	84.81	86.64	87.18

Table 2 SVM parameter optimized by improved FOA

		Iris	Wine	Heart	Vehicle	Waveform
Gauss kernel	γ	1.62	4	0.05	0.249	0.0069
	C	1.84	256	37.7	261	462
Sigmoid kernel	γ	0.20	0.004	0.002	0.009317	0.0051
	c	0	0	0	0	0
	C	36.92	285	56	35861	238
Poly kernel	γ	0.004	0.69	0.01	0.632	4
	c	1	4	4	2	2
	d	3	1	4	3	1
	C	216	1.1	9	8.3	1050
Combination kernel	$\gamma_{rbf} / \gamma_{poly}$	0.5/0.0039	1/0.0625	19.08/0.0318	27.4/0.6	0.005/0.2
	c	0	2	3	1	2
	d	3	5	2	5	2
	C	4	0.5	8.39	1.76	44
	α	0.9	0.8	0.4	0.3	0.7

It can be seen in Table 1 that the classification capability of SVM is quite high after optimized by improved FOA,

To illustrate the significant effect of fixed initial position, the classification accuracy of SVMs when the fruit populations are at initial positions are listed in Table 3.

Table 3 Classification accuracy of SVM at initial position (%)

	Iris	Wine	Heart	Vehicle	Waveform
Gauss kernel	96.67	98.88	83.70	81.21	87.12
Sigmoid kernel	96.67	98.31	83.70	73.05	86.98
Poly kernel	96.67	98.31	84.44	84.99	87.10
Combination kernel	97.33	98.88	84.07	86.05	87.02

As can be seen in Table 3, an initial position close to the optimum value is very important to optimization capability of FOA; it can significantly improve convergence speed of the algorithm, while reduce the computational cost.

In general, kernel function determines the performance of SVM. However, it can be seen in Table 1 and Table 3 that the effect is little for simple data sets and big for complex data sets. For example, the classification accuracy of the sigmoid-SVM based on sigmoid kernel was low in Table 3. Although it performed better after optimized by improved FOA, the classification accuracy is still lower than the other SVMs'.

These experimental results show that the algorithm proposed in this paper is suitable for SVM parameter selection. In addition, it can be found that the method is generally valid for parameter selection of SVM.

Conclusion

The parameter selection is important to SVM research, so far there is no general and effective theory on that problem. Grid search and intelligent algorithm are widely used, but they have some certain shortcomings in application.

An improved FOA is proposed in this paper. The initial position and step size are changed in the algorithm to improve the optimal capability. The algorithm used to optimize the parameter of SVM. The SVMs based on different kernel are used to classify the standard data sets; the effectiveness and generality of the method are proved by the experimental results.

The parameter selection method of SVM based on improved FOA can get the parameters efficiently, it has the advantages of strong generality and simple to use, which make it has a wide application prospect.

Acknowledgements

This paper is supported by National Natural Science Foundation of China (61074127).

References

- [1] B. Baesens, S. Viaene, T.V. Gestel and et al, An Empirical Assessment for Kernel Type Performance for Least Squares Support Vector Machine Classifiers. Brighton, UK: Proceedings of 4th International Conference on Knowledge-based Intelligent Engineering Systems and Allied Technologies. IEEE (2000).
- [2] Y. X. Wu, L. Guo, X. Chai and Y. Wang, On Parameter Selection of Kernel Function Based on Optimisation Algorithm, Computer Applications and Software. Vol. 27 (2010), p. 137- 140.
- [3] X. Y. Yuan and A. L. Liu, Kernel Parameter Selection of the Support Vector Machine Based on Particle Swarm Optimization. Control Theory and Applications. Vol. 26 (2007), p. 5-8.
- [4] W. T. Pan, A New Fruit Fly Optimization Algorithm: Taking the Financial Distress Model as an Example, Knowledge-Based Systems, Vol. 26 (2012), p. 69-74.
- [5] K. T. Fang, M. Q. Liu, Y. D. Zhou. Experimental Design and Modeling, Chemical Industry Press, Beijing (2011).