International Journal for Rapid Research in Engineering Technology & Applied Science
Vol VIII Issue VIII August 2022
ISSN Online: (2455-4723)

Optimizing the parameters of twin bounded support vector machine with genetic algorithm

Shivam Gehlot*, Ass.Prof. Mohit Jain**
Department Of Computer Science Engineering Bm College Of Technology, Indore

---***---

Abstract – *Twin Support Vector machine has bounded on the faster learning speed than classical one. It has attracted many scholars attention through which the sensitivity in the parameters can be selected. In this paper, the parameters selection version of penal and kernel function has set out the quadratic programming optimization. It has been addressed that the application of the Support vector machine (SVM) has presented the improvement and introduction of new optimization of the parameters like FSVM, TSVM, MSVM and other. Furthermore, it has notified that the regularization of the Twin Support Vector machine would help in improving the testing time with genetic algorithm. It can address on the datasets which might be recorded by many scholars to attain accuracy. It has somehow presented that the performance and condition of parallel hyperplanes were relaxed through non-parallel parameters.*

Key Words: Classification, Optimization, Twin bounded Support vector machine, Support vector machines, genetic algorithm.

1. INTRODUCTION

Support Vector Machine (SVM) is one of the finest statistical learning based classification method that maximizes margin classifier. It has set out the geometric margin classifier that provides the two classes in order to reach a minimum generalization error. SVM has managed the diverse classification of the generalized performance. It has critically managed the functional parameters of the kernel functions to optimize on the proper selection along with penalty parameter C. Penalty Parameters has set trade-off between empirical risk and modelling complexity minimization. On the other hand, the Kernel Function parameter has set out the gamma for the radical basis to meet with the non-linear mapping features and input space to some high dimensional feature space.

The Twin Support Vector Regression, novel regression and learning speed obtain a faster yet classical support attention from many scholars. The regression based on cloud particle swarm optimization has characteristics through stable tendency and inertia weight with the basic cloud generation along with the improvement of the diverse population. The Twin Support Vector machine has emerged the machine learning method through which it makes it classified the problems on the proximity and programming problems along with the computational speed as compared to the domains which needs to be promising that applicability. The current research proposes an optimization of the parameters of twin bounded support vector machine with genetic algorithm. This has nested on the two real valued parameters that exist on the techniques where search of all the parameters involves the charge of optimizing it. Therefore, it can reduce the optimization of the magnitude orders in comparison of optimization approaches of traditional method.

2. THEORY OF SUPPORT VECTOR MACHINE

Support Vector Machine Parameters in genetic algorithm provides significant advantages and problems that are separable as well as non-separable. It has included on the section where the theory of SVM would generate three cases.

2.1 Case of Linearly Separable

Through this classification, the situation of linear separable would bring some valuable information regarding prediction or analysis. The algorithms which offered some training data along with corresponding information and result has valuable effects on classification. SVM extends on the two-dimensional way problem to multi-dimensional measures to extend the seed for optimal hyperplane. The optimal hyperplane can be defined as:

$$wT\ x + b = 0, \ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

Suppose a binary classification problem with dataset X ∈ R l×n in which l1 dataset would be positive class and l2 data points are negative class. It has also classified the approximation to what SVMs do need to find a separate line between the two classes dataset.

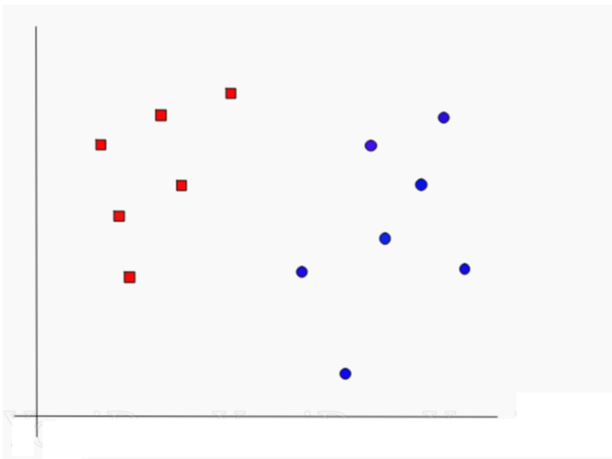

Fig-1: Ideal Line Finding/hyperplane for separate dataset categories

According to the SVM algorithm the points that are close to the both classes lines are support vectors. It can compute on the distances between support vectors and line. It is called margin. This can maximize the goal with the hyperplane and optimal hyperplane features.

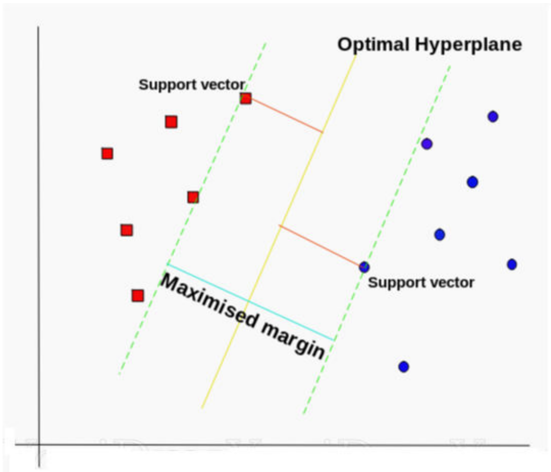

Fig-2 Optimal Hyperplane using SVM algorithm

The decision making boundary in the way to manage the separation between the two classes that are widespread as possible.

2.2 Case of Non-Linearly Separable

The complex data set which is not clear would provide the clarity on the data which can be converted to linearly separable in higher dimensions. This is added with one more dimensions called z-axis. It might coordinate on practices that are governed by constraints,

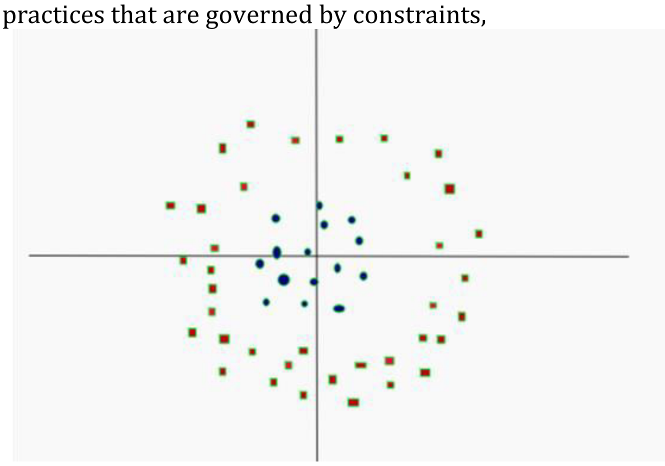

Fig-3: Non-Linearly Separable Data

The equation that represents it as follows:

$$Z= x^2+ y^2 \ldots\ldots\ldots\ldots\ldots (2)$$

The basic coordination for distance square to the point from origin has been plotted on z-axis. It clearly is linearly separable when the data has higher dimensions like z=k, where k is constant. On the other hand, the equation (2) has been equivalent to k then $x^2+ y^2= k$ represents circle equation. This has been linear in meeting original dimensions using transformation parameters. Thus, the data can be classified to extra dimensions and become linearly separable and then project on the boundary decisions with original data. It can certainly focus on kernel function in SVM implementation in the projects.

3. TUNING PARAMETERS: KERNEL, REGULARIZATION, GAMMA AND MARGIN

Kernel

The learning of the hyperplane where SVM needs to transformed on the linear algebra. The kernel parameters manage on the equation for predictions with the new input using the dot product where support vector (xi) is calculated as follows:

F(x)= B(0)+ sum (ai*(x,xi))

The equation involves calculates on the inner products with the input vectors and supporting vector data. The training on the coefficient B0 and estimated with the learning algorithm.

The polynomial kernel can be written with the polynomial and exponential kernel through separation line in higher dimension. This is also called kernel trick.

Regularization

Regularization parameters are often termed as python's sklearn library where SVM optimization needs to be avoided with the classification example. For the larger values of C, optimization needs to meet the smaller margin with the better job application correctly. The classification due to lower regularization value can lead on the higher results like the right one.

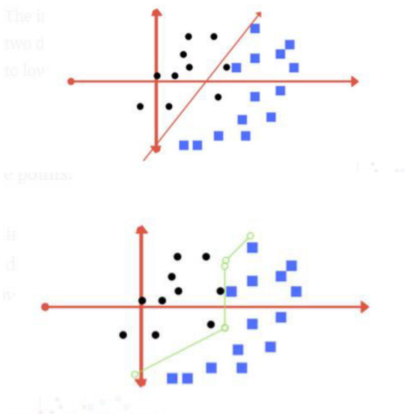

Fig-4: Low regularization value; high regularization value

Gamma

Gamma parameters would define on the training set where the low values means far and high values means close. This can considered as a plausible separation line with the high intense point of the calculation.

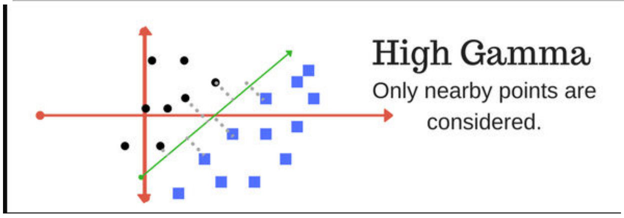

Fig-5: High Gamma

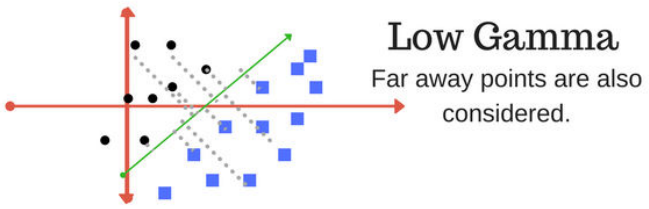

Fig 6: Low Gamma

Margin

The characteristic of SVM classifier would achieve good margin in case of the separation of the classes whether it is good or bad.

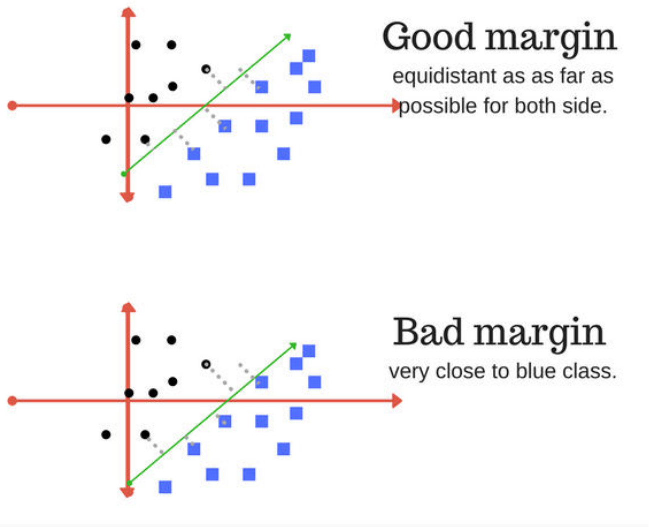

Fig-7: Good Vs Bad margin

SVM tweak on the tuning play with the parameters and implement on the project. The python sklearn has added on the working on support vector machine. High-dimensional spaces on the effective way are to meet the cases where the samples are attained.

The Support Vector Machine is remarkably effective in providing high dimensional spaces as well as setting the versatility. The kernel functions can be specified with the decisions. It can customize kernel and possibly meet the decision function. However, it has featured number of samples which needs to be over-fitting in choosing functions and regularization term in crucial.

4. DISCUSSION

Chou et al (2013) highlights on the hybrid system that deal with the construction engineering and problems of management. The study optimizes on the integrated a fast messy genetic algorithm with support vector machine. The dispute propensity in the initial phase of partnership projects has included measures in term of accuracy, sensitivity, precision and specificity. This has result in providing the pro-active warning and decision-support information needed to manage disputes.

Upadhyay et al (2019) receives the most attention on the machine learning problems and its classification. There is certain algorithm which can be resolved with support vector machine by experts. The SVM has found a decision hyperplane to be in between two supporting parallel where it maximizes the margins as well as improve the accuracy. It has improved the conditions of the parallel hyperplanes along with the regularization of the experiments that are conducted on the real datasets with the model of accuracy. The genetic twin bonded support vector machine has a hybrid GA and TBSVM and algorithm to follow the cross validation techniques can initialize the probabilities of GA parameters.

TWSVM was developed by Jayadeva et al. 2020, for the binary classification issue, and during the past few years, it has become a popular topic for research in machine learning [29]. By resolving two smaller-sized QPPs, such as each hyperplane is as near to one class as possible. The approach of solving two smaller QPPs instead of one larger one is practical. The TSVM's learning rate is accelerated by QPP theoretically, the conventional SVM. Because of TWSVM, It has undergone substantial research and has developed quickly in the last several years due to its decreased computing complexity and improved generalization capability. Variations of the TSVMs, such as twin bounded support, have been proposed.

Theoretically, SVM has a significant benefit in that it can ensure that the extreme solution is the global optimal solution rather than the local minimum. SVM has been successfully applied for classification, regression, and forecasting in real-world settings across all spheres of life. It particularly demonstrates its considerable aptitude for pattern recognition, including text, handwriting, face, etc. SVM is utilized in the realm of economics for stock forecasting, real estate forecasting, etc. SVM is used in the transportation industry to recognize license plate numbers, is trained to differentiate between driving performance and physiological measurement etc. SVM has been utilised in the medical field in automatic medical decision support systems for medical images and forecasting and managing the gene classifications.

5. CONCLUSIONS

SVM is based on statistical learning theory, which methodically investigates the machine learning conundrum, particularly in the context of finite samples. It is based on the structural risk minimum concept and VC-dimensional theory, and by using the kernel function, it can successfully escape the dimensionality curse. These substantial benefits have led to its effective application in numerous fields and good efficiency. Big data has become commonplace along with technological advancement. SVM, however, performs poorly when handling issues involving vast amounts of data. Therefore, much more effort needs to be done to develop SVM further. This includes improving the algorithm's parameter selection and integrating SVM with other disciplines, among other things.

REFERENCES

[1] Avci, E. (2009). Selecting of the optimal feature subset and kernel parameters in digital modulation classification by using hybrid genetic algorithm–support vector machines: HGASVM. *Expert systems with applications*, *36*(2), 1391-1402.

[2] Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, *408*, 189-215.

[3] Chen, K. Y. (2007). Forecasting systems reliability based on support vector regression with genetic algorithms. *Reliability Engineering & System Safety*, *92*(4), 423-432.

[4] Chen, Z., Lin, T., Tang, N., & Xia, X. (2016). A parallel genetic algorithm based feature selection and parameter optimization for support vector machine. *Scientific Programming*, *2016*.

[5] Cheng, L., & Bao, W. (2014). Remote sensing image classification based on optimized support vector machine. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, *12*(2), 1037-1045.

[6] Chou, J. S., Cheng, M. Y., Wu, Y. W., & Pham, A. D. (2014). Optimizing parameters of support vector machine using fast messy genetic algorithm for dispute classification. Expert Systems with Applications, 41(8), 3955-3964.

[7] Ding, S., Zhang, N., Zhang, X., & Wu, F. (2017). Twin support vector machine: theory, algorithm and applications. *Neural Computing and Applications*, *28*(11), 3119-3130.

[8] Fröhlich, H., Chapelle, O., & Schölkopf, B. (2002, December). Feature selection for support vector machines by means of genetic algorithms. In *Proceeding ICTAI* (Vol. 3).

[9] Gupta, P., Mehlawat, M. K., & Mittal, G. (2012). Asset portfolio optimization using support vector machines and real-coded genetic algorithm. *Journal of Global Optimization*, *53*(2), 297-315.

[10] Huang, J., Hu, X., & Yang, F. (2011). Support vector machine with genetic algorithm for machinery fault diagnosis of high voltage circuit breaker. *Measurement*, *44*(6), 1018-1027.

[11] Liao, P., Zhang, X., Li, K., Fu, Y., Wang, M., & Wang, S. (2015). Parameter optimization for support vector machine based on nested genetic algorithms. *Journal of Automation and Control Engineering Vol*, *3*(6).

[12] Liu, H. B., & Jiao, Y. B. (2011). Application of genetic algorithm-support vector machine (GA-SVM) for damage identification of bridge. *International Journal of Computational Intelligence and Applications*, *10*(04), 383-397.

[13] Liu, S., Tai, H., Ding, Q., Li, D., Xu, L., & Wei, Y. (2013). A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction. *Mathematical and Computer Modelling*, *58*(3-4), 458-465.

[14] Omoregbee, H. O., & Heyns, P. S. (2019). Fault classification of low-speed bearings based on support vector machine for regression and genetic algorithms using acoustic emission. *Journal of Vibration Engineering & Technologies*, *7*(5), 455-464.

[15] Sajan, K. S., Kumar, V., & Tyagi, B. (2015). Genetic algorithm based support vector machine for on-line voltage stability monitoring. *International Journal of Electrical Power & Energy Systems*, *73*, 200-208.

[16] Tharwat, A., Hassanien, A. E., & Elnaghi, B. E. (2017). A BA-based algorithm for parameter optimization of support vector machine. *Pattern recognition letters*, *93*, 13-22.

[17] Tomar, D., & Agarwal, S. (2015). Twin support vector machine: a review from 2007 to 2014. Egyptian Informatics Journal, 16(1), 55-69.

[18] Upadhyay, P., Gangrade, J., & Gangrade, S. (2019). Binary Genetic Twin Bounded Support Vector Machine.

[19] Wang, H., Xiong, J., Yao, Z., Lin, M., & Ren, J. (2017, December). Research survey on support vector machine. In *Proceedings of the 10th EAI International Conference on Mobile Multimedia Communications* (pp. 95-103).

[20] Wang, Z., Shao, Y. H., & Wu, T. R. (2013). A GA-based model selection for smooth twin parametric-margin support vector machine. *Pattern Recognition*, *46*(8), 2267-2277.

[21] Wu, C. H., Ken, Y., & Huang, T. (2010). Patent classification system using a new hybrid genetic algorithm support vector machine. *Applied Soft Computing*, *10*(4), 1164-1177.