International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

Optimization of Nonlinear Kernel-Based Classification using Pin-Sgtbsvm

[1]**V Rajanikanth Tatiraju**

[1]Research Scholar, Department of Computer Science and Engineering, P. K. University, Shivpuri, M.P.

[2]**Dr. Rohita Yamaganti**

[2]Associate Professor, Department of Computer Science and Engineering, P. K. University, Shivpuri, M.P.

Abstract

When it comes to machine learning, the selection of kernel functions is crucial for classification model performance. Because they can simulate complicated patterns in data, nonlinear kernel-based classification algorithms have attracted a lot of interest for optimization. Using the German, Haberman, CMC, Fertility, WPBC, Ionosphere, and Live Disorders benchmark datasets as well as others from the UCI database, the paper assesses how well the Pin-SGTBSVM algorithm performs. Using a tenfold cross-validation technique, the ideal parameters are obtained using a nonlinear kernel function. Over six datasets, the findings show that Pin-SGTBSVM outperforms well-known algorithms like TWSVM, TBSVM, Pin-GTWSVM, and Pin-GTBSVM in terms of accuracy, with noticeable advances in classification performance. Although it also shows competitive results, Pin-SGTWSVM's accuracy on the German dataset is marginally worse than TWSVM's. The experimental results show that Pin-SGTBSVM is a reliable method for improving classification accuracy with no impact on computing efficiency. The results show that it might be used for data categorization and machine learning in the actual world.

Keywords: Nonlinear Kernel, Accuracy, Classification, Machine learning, Efficiency

I.INTRODUCTION

Machine learning and data-driven decision-making are dynamic fields, and classification methods are vital for mining large datasets for useful patterns. One of the most effective ways to deal with complicated decision limits that linear classifiers struggle to handle is by using nonlinear kernel-based classification algorithms. In order to translate input data into higher-dimensional feature spaces where linear separation is possible, kernel-based approaches, especially Support Vector Machines (SVM) using kernel functions, offer a way. However, these methods are only as good as the kernel functions, regularization settings, and feature transformation techniques used to select and optimize them. To improve model accuracy, generalizability, and computing efficiency, optimization in nonlinear kernel-based classification is a crucial field of research.

When data distributions display complex interactions that cannot be separated linearly, nonlinear classification difficulties emerge. Due to their linearity assumptions, traditional classifiers like logistic regression and linear discriminant analysis suffer in such cases. In contrast, classifiers that rely on kernel functions do not directly compute the transformation but instead use them to turn input data into feature spaces with greater dimensions. Classifiers can now find computationally feasible, complicated decision boundaries using this method. Some common kernel functions include the sigmoid, polynomial, and Radial Basis Function (RBF), each of which has its own set of benefits that are dependent on the nature of the input. In order to improve model flexibility and fine-tune kernel parameters, robust optimization strategies are required. The selection of kernel function has a substantial influence on classification performance.

Various methodologies are employed in nonlinear kernel-based classification optimization with the goal of enhancing the performance of classifiers. To maximize the balance between model complexity and generalizability, hyperparameter tuning is an essential part. This entails optimizing factors like the regularization coefficient (C) and kernel parameters, such as gamma in RBF kernels. Hyperparameter selection has been automated using grid search, random search, and sophisticated methods like genetic algorithms and Bayesian optimization. Feature selection and dimensionality reduction techniques are also optimized so that classification judgments are based on the most important qualities. It is common practice to use feature scaling and Principal Component Analysis (PCA) to improve

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

input representations and address problems such as high-dimensional sparsity and overfitting.

Computational efficiency is another important optimization metric for kernel-based categorization. The increasing complexity of kernel changes makes training nonlinear classifiers computationally costly, especially with large-scale datasets. Sequential Minimal Optimization (SMO), an effective optimization approach for support vector machine (SVM) training, is one solution that academics have devised to tackle this difficulty. By breaking down large optimization problems into more manageable pieces, these techniques drastically cut down on computing requirements. In addition, we have incorporated parallel and distributed computing frameworks to improve training speed and scalability. This includes GPU acceleration and cloud-based solutions. Random Fourier Features and Nyström approximation are two examples of recent developments in approximate kernel approaches that offer further ways to reduce processing costs without sacrificing classification accuracy.

The use of nonlinear kernel-based classification has many potential uses in many different industries, such as biology, finance, cybersecurity, and image recognition. One use of kernel-based classifiers is in medical diagnostics, where they help with illness prediction and categorization using patient data. This allows for early detection and individualized therapy recommendations. These classifiers find complex patterns in financial transactions and help with credit risk assessment and fraud detection in the financial sector. Object detection and face recognition are two examples of image recognition tasks that greatly benefit from kernel-based classifiers' capacity to distinguish intricate visual characteristics. Additionally, cybersecurity programs protect digital infrastructures from harmful attacks by detecting abnormalities and intrusions in network data using nonlinear classification algorithms. Classifier performance and adaptability should be continuously optimized due to the relevance of these applications.

II.REVIEW OF LITERATURE

Piccialli, Veronica. (2022) Many different areas have found success using support vector machines, making them a crucial class of machine learning models and techniques. To define the machine learning models and to create effective and convergent algorithms for large-scale training tasks, SVM technique relies heavily on nonlinear optimization. Here, we lay out the convex programming issues that underpin support vector machines (SVMs), with an emphasis on supervised binary classification. Here, we take a look at the most popular optimization techniques for support vector machine (SVM) training issues and talk about how to incorporate their characteristics into algorithm design.

Li, Kai & Lv, Zhen. (2021) To improve the support vector machine's classification performance, the twin support vector machine resolves two small quadratic programming problems. The following problems, however, afflict this method: (1) The twin support vector machine and other variants rely on a hinge loss function to construct their models, however this function is noisy and unstable during resampling. (2) Getting the models to work in the dual space takes a lot of time and effort. To make the twin bounded support vector machine even more effective, the pinball loss function is included into it. To solve the problem of the pinball loss function not being differentiable at zero, a smooth approximation function is built. One may use this to build a smooth twin-bounded support vector machine model that includes pinball loss. Iteratively, the issue is solved in the original space using the Newton-Armijo approach. Theoretically, we show that an iterative method for smooth twin bounded support vector machines with pinball loss converges. The trials verify the suggested approach on both real and synthetic datasets, including those from UCI. In addition, the suggested algorithm's efficacy is shown by comparing its performance to that of other representative algorithms.

Yao, Yukai et al., (2015) We present PMSVM, an enhanced Support Vector Machine classifier that takes into account extensively System Normalization, PCA, and Multilevel Grid Search techniques for data pretreatment and parameters optimization, respectively. Improving SVM's classification efficiency and accuracy are the primary objectives of this project. To evaluate PMSVM's efficacy, metrics such as ROC curve, sensitivity, specificity, and precision are utilized. When compared to more conventional SVM algorithms, experimental findings reveal that PMSVM significantly outperforms them in terms of efficiency and accuracy.

Cocianu, Catalina. (2013) This research details a model-free method for developing SVM-type non-linear classifiers. In common parlance, support vector machines are "non-parametric" models. However, this does not mean that SVMs do not have parameters; rather, the learning issue around the parameters of an SVM is of paramount significance. A revised version of the gradient ascent method for addressing the SVM - QP issue and a new formulation for the bias parameter provide the innovative aspects of the suggested approach. In addition to demonstrating greater convergence rates than the basic SMO method, the tests also highlighted good convergence qualities of the suggested modified variations. In comparison to the default bias setting, the produced classifier also performs better.

Biswas, Debojit et al., (2011) For a wide variety of reasons, land cover data is crucial. Precise data on a region's land cover

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 November 2023*

is essential for many initiatives that aim to manage, plan, and monitor natural resources. An image classification technique is used to recover land cover information from remotely sensed photos, which are interesting sources for this purpose. When training data is scarce and classes have few pixels, statistical classifiers often fail. Support vector machines (SVMs) and other next generation algorithms have been producing respectable results with a less quantity of training data sets. Training is an ongoing expense. As a result, there is a great deal of interest in developing classifiers that require a smaller set of training data. We examine support vector machine (SVM) based hyperspectral image classification using several kernel types, including linear, polynomial, radial basis, and sigmoid, in this research. We test SVM's performance with various kernels and compare the results. Additionally, this section delves into the mathematical foundations of non-linear SVM. For the purpose of feature reduction, this study used principal component analysis (PCA). Since the penalty amount has less of an impact on linear and polynomial kernels in SVM, our results demonstrate that these kernel types achieve better classification accuracy. A narrow range of penalty levels and hyperparameters is required for other kernels.

III.MATERIAL AND METHODS

Using the UCI database, we conduct trials on the following categories: German, Haberman, CMC, Fertility, WPBC, Ionosphere, and Live-disorders to validate the performance of the Pin-SGTBSVM. In this experiment, the kernel function is $K\,(x,y)\;=\;exp(-\theta\times\|x-y\|^2)$, where θ is a parameter, and the best parameter value within the range is determined using a tenfold cross-validation procedure [10^{-6},10_5]. The values of τ1 and τ2 are 0.5, 0.8, 1, and ci> 0 (i = 1, 2,3,4) and the value range is [2^{-10}, 2^{10}], the value of ε in the experiments is 10_{-6}, the value of η in the algorithm is 10^{-4}, and the standard deviation and average accuracy are included in the experimental findings. An Intel (R) Core (TM) i7-5500U CPU running at 2.40GHz with 4GB of RAM and MATLAB R2016a were utilized for all the experiments presented in this paper.

We compare the performance of many representative methods, including TWSVM, TBSVM, Pin-GTWSVM (TBSVM with pinball loss in dual space), and Pin-GTBSVM (TBSVM with pinball loss in original space). The iterative approach is also used to solve these algorithms.

IV.RESULTS AND DISCUSSION

Results from the experiments are displayed in table 1. By comparing the six datasets, it is clear that the Pin-SGTBSVM algorithm outperforms the other five approaches. However, when it comes to the Fertility dataset, the Pin- SGTBSVM methodology achieves identical results to the other five.

Additionally, whereas the TWSVM method has superior accuracy on the German dataset, the Pin-SGTWSVM approach achieves greater accuracy on all six datasets.

Table 1 Comparison of Algorithm Performance Using Nonlinear Kernels

Datasets	**TWSVM**	**TBSVM**	**Pin-GTWSVM**	**Pin-GTBSVM**	**Pin-SGTWSVM**	**Pin-SGTBSVM**
	Accuracy (%)	**Accuracy (%)**	**Accuracy (%)**	**Accuracy (%)**	**Accuracy (%)**	**Accuracy (%)**
	±sd	**±sd**	**±sd**	**±sd**	**±sd**	**±sd**
	Time(s)	**Time(s)**	**Time(s)**	**Time(s)**	**Time(s)**	**Time(s)**
German	74.9170	70.0999	70.0999	70.0996	71.3955	75.1497
	±0.0189	±0.0002	±0.0002	±0.0002	±0.0262	±0.018
	0.0875	0.0563	0.0865	0.0742	0.0304	0.0375
Haberman	73.3698	73.3694	73.3697	73.3697	73.7498	73.7497
	±0.0001	±0.0001	±0.0001	±0.0001	±0.0062	±0.0062
	0.0294	0.0326	0.0280	0.0309	0.0240	0.0184
CMC	69.2387	65.396	65.3980	65.3980	71.662	73.7026
	±0.0398	±0.0000	±0.0000	±0.0000	±0.0359	±0.0227
	0.0695	0.0570	0.0635	0.0497	0.0173	0.0405
Fertility	87.0971	87.0967	87.0966	87.0966	87.0970	87.0971
	±0.0002	±0.0002	±0.0002	±0.0000	±0.0000	±0.0002
	0.0144	0.0121	0.0141	0.0147	0.0151	0.0108
WPBC	76.2713	76.2715	76.2713	76.273	78.8134	79.1526
	±0.0000	±0.0001	±0.0000	±0.0000	±0.0349	±0.0429
	0.0148	0.0197	0.0295	0.0294	0.0197	0.0211

Ionosphere	91.6980	91.6980	93.5850	92.7360	93.1130	94.5285
	±0.0252	±0.0130	±0.0328	±0.0298	±0.0215	±0.0142
	0.0252	0.0247	0.0349	0.0371	0.0174	0.0263
Live disorders	64.4229	62.6925	57.6926	57.6925	66.8272	67.1156
	±0.0495	±0.0460	±0.0000	±0.0001	±0.0480	±0.0577
	0.0638	0.0593	0.0507	0.0598	0.0161	0.0172

V.CONCLUSION

The accuracy of the Pin-SGTBSVM algorithm was found to be higher than that of other approaches, including TWSVM, TBSVM, Pin-GTWSVM, and Pin-GTBSVM, when tested on many datasets from the UCI repository. Across all datasets, but notably German, CMC, WPBC, and Ionosphere, the algorithm demonstrated superior performance. It is worth mentioning that Pin-SGTBSVM performed similarly to other methods on the Fertility dataset. Additionally, Pin-SGTWSVM accomplished respectable accuracy; the only dataset where it was somewhat less accurate than TWSVM was the German one. In particular, the experimental findings show that Pin-SGTBSVM performs well over a wide range of real-world datasets, with respect to accuracy, standard deviation, and computing efficiency. Based on these results, the Pin-SGTBSVM method seems like it may be a great tool for accuracy and computing efficiency classification problems in many different disciplines.

REFERENCES: -

[1] V. Piccialli, "Nonlinear optimization and support vector machines," *Annals of Operations Research*, vol. 314, no. 1, pp. 111–149, 2022.

[2] K. Li and Z. Lv, "Smooth twin bounded support vector machine with pinball loss," *Applied Intelligence*, vol. 51, no. 1, pp. 1–17, 2021, doi: 10.1007/s10489-020-02085-5.

[3] F. Liu, X. Huang, Y. Chen, et al., "Random features for kernel approximation: A survey in algorithms, theory, and beyond," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 7128–7148, 2021.

[4] B. Jiang, T. Lin, S. Ma, et al., "Structured nonconvex and nonsmooth optimization: Algorithms and iteration complexity analysis," *Computational Optimization and Applications*, vol. 72, no. 1, pp. 115–157, 2019.

[5] Y. L. Feng, Y. N. Yang, X. L. Huang, S. Mehrkanoon, and J. A. K. Suykens, "Robust support vector machines for classification with nonconvex and smooth losses," *Neural Computation*, vol. 28, no. 6, pp. 1217–1247, 2016.

[6] Y. Yao, H. Cui, Y. Liu, L. Li, L. Zhang, and X. Chen, "PMSVM: An optimized support vector machine classification algorithm based on PCA and multilevel grid search methods," *Mathematical Problems in Engineering*, vol. 2015, no. 1, pp. 1–15, 2015, doi: 10.1155/2015/320186.

[7] C. Cocianu, "Kernel-based methods for learning non-linear SVM," *Economic Computation and Economic Cybernetics Studies and Research / Academy of Economic Studies*, vol. 47, no. 1, pp. 1–15, 2013.

[8] D. Biswas, H. Jain, M. Arora, and B. Balasubramanian, "Study and implementation of a non-linear support vector machine classifier," *International Journal of Earth Sciences and Engineering*, vol. 4, no. 0974-5904, pp. 338–341, 2011.

[9] K. Huang, R. Zheng, R. Sun, R. Hotta, and R. Fujimoto, "Sparse learning for support vector classification," *Pattern Recognition Letters*, vol. 31, no. 13, pp. 1944–1951, 2010.

[10] T. Hofmann, B. Scholkopf, and A. J. Smola, "Kernel methods in machine learning," *Annals of Statistics*, vol. 36, no. 3, pp. 1171–1220, 2008.

[11] C. Orsenigo and C. Vercellis, "Multivariate classification trees based on minimum features discrete support vector machines," *IMA Journal of Management Mathematics*, vol. 14, no. 2, pp. 221–234, 2003.

[12] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing multiple parameters for support vector machines," *Machine Learning*, vol. 46, no. 1-3, 2002.