Issues in Information Systems
Volume 27, Issue 4, pp. 209-223, 2026

DOI: https://doi.org/10.48009/4_iis_2026_117

Analytical parameter selection with critical computational resources

Manuel Ramírez, *Universidad Federal de Paraíba, authoremail@email.com*
José Luyo, *Universidad Nacional Mayor de San Marcos, jluyos@unmsm.edu.pe*
Ana Luna, *Universidad del Pacífico, ae.lunaa@up.edu.pe*
Lucas Machuca, *Universidad del Pacífico, lg.machucag@up.edu.pe*
Mario Chong, *Universidad del Pacífico, m.chong@up.edu.pe*

Abstract

Support Vector Machines (SVMs) are essential for machine learning classification, yet their performance depends heavily on Gaussian Kernel hyperparameter selection. Traditional iterative tuning methods, such as cross-validation and heuristics, suffer from high computational complexity, limiting real-time application under constrained resources. This study addresses these limitations by proposing a geometric-differentiable framework for analytical parameter selection based on class separability. By incorporating direct and inverse variation criteria governed by a non-linear control exponent, the proposed approach allows instantaneous, closed-form adjustments of optimization weights. Algorithmic scalability was validated using the large-scale ijcnn1 benchmark, achieving a 99.8% reduction in hyperparameter selection time compared to conventional Grid Search. Additionally, the Statlog (German Credit Data) dataset was utilized as an analytical proxy to simulate socioeconomic capacity assessments within post-crisis environments. Results show that the proposed analytical criteria maintain high classification accuracy, reaching up to 98.8% in testing sets. This mathematical optimization resolves the conflict between computational speed and operational flexibility, establishing a plausible, highly efficient framework to enhance decision-support systems when processing power is critically limited.

Keywords: Support Vector Machine, Gaussian Kernel, Humanitarian, Logistics.

Introduction

In contemporary data-driven applications, high-dimensional datasets have become essential for modeling complex phenomena and extracting relevant patterns from large volumes of information. However, the continuous growth of available data has introduced significant challenges for real-time decision-making, as conventional analytical approaches may be insufficient to capture complex relationships efficiently. Machine learning methods provide a systematic alternative by enabling automated pattern recognition and predictive modeling through mathematical optimization frameworks. Among these techniques, the Support Vector Machine (SVM) framework, initially introduced in its linear formulation by Vapnik and Chervonenkis (1964), remains a widely adopted approach for binary and multiclass classification tasks due to its foundations in statistical learning theory and structural risk minimization (Vapnik, 1995).

The performance of non-linear SVM models relies on kernel functions, which enable the representation of input data in higher-dimensional feature spaces according to the principles of Reproducing Kernel Hilbert Spaces (Aronszajn, 1950; Mercer, 1909). In particular, the Gaussian Radial Basis Function (RBF) kernel is frequently employed due to its topological advantages, such as structure-preservation (Wang et al., 2003)

and its high-dimensional Kernel space (Cover, 1965). However, the predictive performance of RBF-based SVM classifiers is highly dependent on the appropriate selection of the kernel width parameter, commonly denoted as σ. This parameter determines the influence range of individual samples and directly affects the complexity of the resulting decision boundary. Therefore, obtaining an adequate value of σ is a fundamental requirement for achieving a suitable balance between classification accuracy and generalization capability.

Common searching methods focused on identifying this hyperparameter primarily rely on iterative strategies, such as Cross-validation (Hsu et al., 2003) or heuristic searching algorithms (Li et al., 2010), which often suffer from high computational time complexity drawbacks. To address these limitations, Liu et al. (2015) proposed an analytical, closed-form method based on the concept of class separability, which represents the similarity between pairs of points to estimate the kernel parameter through an objective function. Although this approach significantly reduces the time required for hyperparameter configuration, it relies on a fixed couple of penalization parameters, offering no mechanism to adjust or control them according to the specific requirements of different classification problems. This mathematical rigidity restricts the adaptability of the classifier, as it prevents the model from shifting its optimization focus when facing distinct operational priorities.

This limitation becomes highly critical in high-pressure scenarios such as disaster management and humanitarian logistics. In the immediate aftermath of a natural crisis, relief agencies and governments must deploy rapid decision-support systems to assess the socio-economic capacities of affected populations and distribute limited emergency subsidies efficiently. While iterative optimization methods like Cross-Validation are operationally impractical in field-level crisis zones due to compromised power grids and legacy hardware, a rigid analytical model is also insufficient. Field environments demand a flexible framework that can be dynamically adjusted, allowing decision-makers to prioritize either within-class compactness to prevent resource misallocation or between-class diversity to ensure that no vulnerable individual is left without aid.

The current work addresses this gap by proposing two analytical methods for penalization parameter selection based on direct and inverse variation criteria. By introducing a non-linear control exponent, the proposed formulations extend the mathematical scope of class-separability frameworks, enabling instantaneous, closed-form adjustments of the optimization weights. This approach resolves the conflict between computational speed and operational flexibility, establishing a plausible mathematical framework to enhance decision-support systems even when computational resources are critically constrained.

Literature review

The SVM has an outstanding theoretical background to deal with classification tasks due to being founded upon the idea of risk minimization (Vapnik, 1995). In practical industrial applications, SVM classifiers are widely utilized for machinery condition monitoring, structural fault diagnosis, and engineering reliability management, where classification accuracy directly influences operational safety and resource allocation efficiency (Widodo & Yang, 2007). Beyond heavy industry, the algorithm serves as a foundational tool in computational biology for high-dimensional gene expression profiling and text categorization, as well as in financial engineering for credit scoring and bankruptcy prediction (Huang et al., 2007). By identifying optimal hyperplanes in sparse or complex vector spaces, SVMs effectively model intricate non-linear relationships within sensitive socioeconomic datasets. More recently, this predictive capability has been extended to humanitarian logistics and disaster management scenarios, where rapid vulnerability screening is essential for equitable post-crisis reconstruction and resource deployment.

Support Vector Machines for Classification

Roughly speaking, the goal of SVM's, for binary classification, is to construct a hyperplane that separates the space where a training set of examples is given, so that the elements which belong to the same class stay in the same half-space and that new surface works well on unseen cases.

In order to construct a SVM, firstly, the training set shall be considered $T = \{(\vec{x}_1, y_1), \ldots, (\vec{x}_l, y_l)\} \subset X \times \{-1, 1\}$, where X ⊂ Rn if this set is linearly separable, which means that at least a separating hyperplane that splits one class from the other, exists; this problem is treated by a hard margin support vector machine approach which searches for a hyperplane that separates the training set by means of maximizing the distance between the two closest classes of vectors and the hyperplane, which in turn is defined by an orthogonal vector $\vec{\omega} \in R^n$ to the hyperplane and an offset term b ∈ R. Thus, the decision function can be defined by $f(x) = sgn(\langle \vec{w}, \vec{x} \rangle + b)$, where sgn turns out 1 if its argument is greater than or equal to zero or otherwise, -1. The optimization problem for hard margin SVM is as follows:

$$minimize_{\vec{\omega},b} : \frac{1}{2}\|\vec{\omega}\|^2$$
$$\text{s.t.} \quad (\langle \vec{\omega}, \vec{x}_i \rangle + b) y_i \geq 1; \; i = 1,\ldots,l \tag{1}$$

Considering the dual problem of the question (1), if $\vec{\alpha}^* = (\alpha_1^*,\ldots,\alpha_l^*)$ is a solution to it, then there must be a nonzero component αj∗ (Deng et al., 2013) and the solution to the problem (1) is obtained in the following way:

$$\vec{\omega}^* = \sum_{i=1}^{l} \alpha_j^* y_i x_i \tag{2}$$

$$b^* = y_j - \sum_{i=1}^{l} \alpha_j^* y_i \langle \vec{x}_i, \vec{x}_j \rangle \tag{3}$$

When data is not separable, the strategy is to map the data $\{\vec{x}_1,\ldots,\vec{x}_l\} \in X$ into a higher-dimensional Hilbert space F where the data will probably be linearly separable (Cover, 1965), which is called feature space, via a function ϕ : X → F named the feature mapping. Therefore, there is a relation between instances in the input space X and feature space F by a Kernel function $K : X \times X \to R$, which satisfies (Berlinet & Thomas-Agnan, 2004):

$$K(\vec{x}, \vec{y}) = \langle \phi(\vec{x}), \phi(y) \rangle_F \tag{4}$$

Thus to solve such non-separable classification problem, we could consider the same algorithm in F instead of X, i.e., and we could work with the training set $\{(\phi(\vec{x}_1), y_1),\ldots,(\phi(\vec{x}_l), y_l)\} \subset F \times \{-1, 1\}$, which is as known as the Kernel trick (Herbrich, 2002).

However, in most of the classification problems, such a training set in the feature space F may still not be separable, so the Soft Margin Support vector Machine must be used in this case by introducing slack variables ξ1,...,ξl, into the equation (1), hence the optimization problem of Soft Margin SVM turns out to be:

$$minimize_{\omega,b,\vec{\xi}} : \frac{1}{2}\|\omega\|^2 + C\sum\nolimits_{i=1}^{l} \xi_i$$
$$\text{s.t.} \quad (\langle \omega, \phi(\vec{x}_i) \rangle + b) y_i \geq 1 - \xi_i; \; \omega \in F; \; \xi_i \geq 0; \; \text{i} = 1,\ldots,l \tag{5}$$

To solve the problem (5), the Lagrange multiplier method (Griva et al., 2009) is used:

$$maximize_{\lambda}: -\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\lambda_i\lambda_j y_i y_j \langle \phi(\vec{x}_i), \phi(\vec{x}_j)\rangle_F + \sum_{i=1}^{l}\lambda_i$$

$$\text{s.t. } \sum_{i=1}^{l} y_i\lambda_i = 0 \tag{6}$$

$$0 \le \lambda_i \le C;\ \vec{\lambda} \in R^l;\ \text{i=1,....,}l$$

Furthermore, if $\vec{\lambda}^*=(\lambda_1^*,...,\lambda_l^*)$ is a solution to the problem (6), then for any component $\vec{\lambda}_j^*\ of\ \vec{\lambda}^*$ such a $\lambda_j^* \in (0, C)$, the solution (ω∗, b∗) to the problem (5) can be obtained (Deng et al., 2013) as follows:

$$\omega^* = \sum_{i=1}^{l}\lambda_i^* y_i \phi(\vec{x}_i) \tag{7}$$

$$b^* = y_j - \sum_{i=1}^{l}\lambda_i^* y_i \langle \phi(\vec{x}_i), \phi(\vec{x}_j)\rangle_F = y_j - \sum_{i=1}^{l}\lambda_i^* y_i K(\vec{x}_i, \vec{x}_j) \tag{8}$$

where the relation (4) has been used in the last inequality. Therefore (4) and (7) lead to the decision function:

$$f(\vec{x}) = sgn(\langle \omega^*, \phi(\vec{x})\rangle_F + b^*) = sgn(\sum_{i=1}^{l}\lambda_i^* y_i K(\vec{x}_i, \vec{y}) + b^*), for\ all\ \vec{x}\ \in\ R^n \tag{9}$$

Any Mercer's Kernel (Schölkopf & Smola, 2002) can be used as a Kernel function. Among the Kernel functions that can be chosen, the polynomial kernel $K(\vec{x}, \vec{y}) = (\langle \vec{x}, \vec{y}\rangle + c)^p, c \in R^+, p \in N$, and the Gaussian kernel $K(\vec{x}, \vec{y}) = exp(\|\vec{x} - \vec{y}\|/2\sigma^2), \sigma \in R^+$, are common examples.

As well as dealing with binary classification tasks, the SVM can be extended to multiclass classification, namely for training sets like it is in Eq. 10:

$$T_M\ \{(\vec{x}_1,y_1),...,(\vec{x}_l,y_l)\} \subset R^n \times Y_m \tag{10}$$

where Ym = {1,2,3,...,M} for some M ∈ N. There are many methods to deal with multiclass classification (Tsoumakas & Katakis, 2007) and most of them are based on transforming such a problem into a set of binary classification problems (Hsu & Lin, 2002).

Training Support Vector Machines

The training process of a Support Vector Machine requires the appropriate selection of a set of hyperparameters that determine the behavior of the resulting classifier. Among these parameters, the soft margin coefficient *C* from Equation 5 and the internal parameters of the selected kernel function are the most relevant. Their configuration directly influences the trade-off between model complexity and generalization capability. A large value of *C* increases the penalty assigned to classification errors, encouraging the optimization process to correctly classify training samples and producing narrower decision margins. However, excessively large values may lead to overly complex decision boundaries that capture noise or specific characteristics of the training data, reducing the model's ability to generalize to unseen observations. Conversely, smaller values of *C* allow greater tolerance to misclassified samples and produce wider margins, but excessively low values may result in insufficient model complexity and underfitting, preventing the classifier from adequately representing the underlying data distribution (Hsu et al., 2003).

For non-linear classification problems, the Gaussian Radial Basis Function (RBF) kernel is one of the most adopted alternatives due to its ability to map input samples into a higher-dimensional feature space. However, its performance is strongly influenced by the kernel width parameter (σ), which determines the degree of locality of the resulting similarity measure and consequently affects the shape of the decision boundary. When σ assumes very small values, the kernel becomes highly localized, allowing the classifier to generate complex boundaries that closely follow individual training samples. Although this configuration can improve training accuracy, it may reduce generalization capability by increasing sensitivity to noise and sample-specific variations. In contrast, very large values of σ produce smoother decision boundaries because the influence of individual samples extends over a wider region of the feature space. Consequently, the separation between classes may become insufficient, leading to underfitting and reduced predictive performance (Wang et al., 2003).

The strong dependency of SVM performance on the joint configuration of *C* and σ creates a challenging hyperparameter optimization problem, where only a limited region of the parameter space generally provides an appropriate balance between classification accuracy and model complexity. To identify this optimal configuration, several approaches have been proposed in the literature. The most widely adopted strategy is Grid Search combined with cross-validation (Hsu et al., 2003), which evaluates multiple combinations of hyperparameters through repeated model training. Although effective, this approach becomes computationally expensive as the search space increases, since every candidate configuration requires a complete optimization process. To reduce this computational burden, alternative methods based on heuristic optimization have been explored, including genetic algorithms and gradient-based strategies (Chapelle et al., 2002; Zhao et al., 2011). These approaches can improve the search efficiency but still depend on iterative procedures that may become impractical in time-constrained applications. Consequently, recent research has focused on analytical and non-iterative parameter selection methods, including the class-separability approach proposed by Li et al. (2010) and the closed-form framework introduced by Liu et al. (2015), which estimate suitable kernel parameters directly from statistical characteristics of the dataset without requiring exhaustive iterative optimization.

Research Methodology

Zhiliang's Method to Select the Gaussian Parameter

Let us consider a training set S = $\{(\vec{x}_1,y_1),...,(\vec{x}_N,y_N)\} \subseteq Rn \times \{1,2,...,L\}$ where the element $x_i^{(k)} \in S$ denotes the i-th element belonging to the k-th class and each class i has Ni elements, i = 1,…,L. Zhiliang proposes to use the class separability concept for describing how data is scattered in the feature space by considering two principles.

- Principle I: Instances from the same class should be as similar as possible and
- Principle II: Instances from different classes should be as diverse as possible.

Considering the Gaussian Kernel and Normal distributed data, there is a relationship between distances in the input space X ⊂ Rn and the feature space F and it is given by,

$$\|\phi(\vec{x}) - \phi(\vec{y})\|^2 = 2 - 2e^{-\|\vec{x}-\vec{y}\|^2/2\sigma^2} \quad (11)$$

for $\vec{x}, \vec{y} \in R^n$, where ϕ: Rn → F is the feature mapping. Thus, we define W' and B' by

$$W' = \frac{1}{\sum_{r=1}^{L} N_r^2} \sum_{i=1}^{L} \sum_{t=1}^{L} \sum_{k=1}^{L} \left\| \overrightarrow{x_t^{(i)}} - \overrightarrow{x_k^{(j)}} \right\|^2 \quad (12)$$

$$B' = \frac{1}{\sum_{r=1}^{L}\sum_{s=1}^{L} N_r N_s} \sum_{i=1}^{L}\sum_{j=i+1}^{L}\sum_{t=1}^{L}\sum_{k=1}^{L} \left\| \overrightarrow{x_t^{(i)}} - \overrightarrow{x_k^{(j)}} \right\|^2 \quad (13)$$

where W′ and B′ are the average of training points of the same class and different classes respectively. As an extra assumption, B′ > W′ must be satisfied to proceed with this method. According to (11), (12), (13) and the assumption of normally distributed data, it is possible to define within-class and between-class separability given by W and B respectively as follows,

$$W = 2 - 2e^{-\frac{W'}{2\sigma^2}} \quad (14)$$

$$B = 2 - 2e^{-\frac{B'}{2\sigma^2}} \quad (15)$$

Based on the announced separability principles previously mentioned, Zhiliang formulates the function of class separability in the following way

$$\textit{maximize } J(\sigma) = \omega_B B(\sigma) - \omega_W W(\sigma), \quad \text{s.t. } \sigma > 0 \quad (16)$$

where ω_W, $\omega_B > 0$ are penalization parameters. Analyzing this function, it has a unique solution:

$$\sigma^* = \sqrt{\frac{B' - W'}{2ln(\omega_B B'/\omega_W W')}} \quad (17)$$

which exists whenever

$$\omega_B B' > \omega_W W' \quad (18)$$

It can be noticed that if $\omega_B < \omega_W$, the function J will focus on maximizing the term B and hence lying elements belonging to the same class as close as possible and if $\omega_B > \omega_W$, J will lie elements from different classes as far as possible. Finally, Zhiliang proposes two values for penalization parameters: $\omega_W = \omega_B = 0.5$ or $\omega_W = \frac{W'}{W'+B'}, \omega_B = \frac{B'}{W'+B'}$, hence, those values can be replaced in the equation (17) to get the optimal parameter of the Gaussian Kernel.

Proposed Methods for Penalization Parameters Selection

The penalization parameters ω_W and ω_B in the aggregate objective function $J(\sigma) = \omega_B B(\sigma) - \omega_W W(\sigma)$, play a critical role in regularizing the geometric structure of the kernel space. To understand how managing the trade-off between within-class similarity $W(\sigma)$ and between-class diversity $B(\sigma)$ operates in a real-world decision support system, we can map these mathematical components directly onto the German Credit Data framework adapted for humanitarian logistics.

The dataset models a critical post-disaster scenario where relief agencies must classify a population of size l into two categories: individuals requiring urgent financial subsidies for reconstruction ($y_i = -1$) and self-sufficient individuals ($y_i = 1$). The mathematical adjustment of the weights directly alters the operational risk management of resource distribution as follows:

- Focus on Within-Class Compactness ($\omega_B < \omega_W$): When the analyst penalizes within-class variance more heavily, the optimization of $J(\sigma)$ prioritizes minimizing $W(\sigma)$, forcing instances of the same class to lie as close as possible in the feature space. Logistically, this forces a highly cohesive and tight geometric definition of the 'vulnerable profile.' This setting is desirable when computational and physical resources are critically constrained, as it minimizes the geographic dispersion of the target boundary to prevent aid allocation errors (aid fraud).

- Focus on Between-Class Diversity ($\omega_B > \omega_W$): Conversely, prioritizing between-class diversity forces the solution σ^* to maximize $B(\sigma)$, pushing different types of instances as far apart as possible. In a crisis zone, this configuration maximizes the distance between the vulnerable population and self-sufficient individuals. Consequently, this minimizes the risk of false negatives near the boundary, ensuring that no eligible victim is accidentally left without support due to data overlap.

To mathematically operationalize this trade-off without reverting to high-cost iterative algorithms, we propose two distinct non-linear criteria governed by a control exponent $p \geq 1$. The first, based on an inverse-variation criterion, forces a dynamic equilibrium ($\omega_W^p + \omega_B^p = 1$) to prevent scale dominance. The second, based on a direct-variation criterion, establishes a proportional growth trajectory ($\omega_B = \omega_W^p$). In both approaches, the exponent p acts as an exact mathematical knob that controls boundary sensitivity, converting a multi-hour computational optimization into an instantaneous, closed-form algebraic evaluation.

Selection Method Based on Inverse-Variation Criterion

Selection method based on Inverse-variation criterion

The following conditions lead to choosing the penalization parameters ω_W y ω_B for (16) and are given by

1. $\omega_W > 0$, $\omega_B > 0$,
2. $\omega_B B' > \omega_W W'$,
3. $\omega_W^p + \omega_B^p = 1,\ p \geq 1.$

Condition 1 arises from the fact that ω_W and ω_B are penalization parameters, hence they must be positive. Condition 2 guarantees the existence of the solution (17) and the last condition controls possible long differences between the terms in (16) while one parameter grows the other one decreases its value and vice-versa. This fact gives rise to the name of the current method proposed. The parameter $p \in R$ in the third condition takes control of the degree of the difference between such two penalization parameters.

$$\omega_B B' > \omega_W W' \tag{18}$$

It can be proved that the three conditions above produce the following permissible range of values for the penalization parameters:

$$0 < \omega_B < \frac{B'}{\sqrt[p]{B'^p + W'^p}}, \quad \forall p > 1 \tag{19}$$

and

$$\frac{W'}{\sqrt[p]{B'^p + W'^p}} < \omega_B < 1, \quad \forall p > 1 \tag{20}$$

To supply two penalization parameters who satisfy those ranges, it can be noticed that $W'/\sqrt[p]{B'^p + W'^p} < B'/\sqrt[p]{B'^p + W'^p}$ since $W' < B'$, so for a $p \geq 1$ given, ω_B can be considered as

$$\omega_B^* = \frac{B' + W'}{2\sqrt[p]{B'^p + W'^p}} \tag{21}$$

and the value for ω_W is given by

$$\omega_W^* = \sqrt[p]{1-(\omega_B^*)^p} = \sqrt[p]{1-\frac{(B'+W')}{2^p(B'^p+W'^p)}} \tag{22}$$

Note that ω_B is just the average of $W'/\sqrt[p]{B'^p+W'^p}$ and $B'/\sqrt[p]{B'^p+W'^p}$. It can also be proved that those values satisfy

$$\omega_W^* < \omega_B^*, \quad \forall p > 1 \tag{23}$$

which is an important fact if we want to focus the SVM training on having instances as close as possible of the same.

A similar procedure can generate an analogous couple of penalization parameters to get instances belonging to different classes as far as possible. For that purpose, the values of penalization parameters can be regarded as follows:

$$\omega_W^{**} = \frac{B'+W'}{2\sqrt[p]{B'^p+W'^p}} \tag{24}$$

and

$$\omega_B^{**} = \sqrt[p]{1-(\omega_B^{**})^p} = \sqrt[p]{1-\frac{(B'+W')}{2^p(B'^p+W'^p)}} \tag{25}$$

which in turn satisfies the desired aim:

$$\omega_W^{**} < \omega_B^{**}, \quad \forall p > 1 \tag{26}$$

Selection method based on Direct-variation criterion

The previous method was based on inverse-variation criterion for the penalization parameters. Now, we introduce another method focused on direct variation on penalization parameters. To do so, we consider the well-known conditions:

- $\omega_W > 0$, $\omega_B > 0$ (for being parameters of punishment),
- $\omega_B B' > \omega_W W'$ (for the existence of the σ^*solution).

We propose two conditions focused on the within-class separability ($\omega_B < \omega_W$) and the between-class separability ($\omega_B > \omega_W$) in the same way as in the former method. First, to focus on within-class separability, two more conditions are proposed:

I. $\omega_W < 1$,
II. $\omega_B = \omega_W^p$, $p > 1$.

The first condition ensures that penalization parameters will not alter the terms in the function (16) extensively, and last condition controls a proportional increase or decrease in the penalization parameters. As in the previous method, conditions i. and ii. lead to a permissible range for the penalization parameters:

$$\left(\frac{W'}{B'}\right)^{\frac{p}{p-1}} < \omega_W < 1, \quad \forall p > 1 \tag{27}$$

and

$$\left(\frac{W'}{B'}\right)^{\frac{p}{p-1}} < \omega_B < 1, \quad \forall p > 1 \tag{28}$$

Furthermore, in order to get two values for penalization parameters which satisfy the previous ranges, we can consider for any $p > 1$, the value of ω_B given by

$$\omega'_B = (\frac{W'}{B'})^{\frac{p+1}{2p(p-1)}} \quad (29)$$

and ω_W we may assume the value produced by the condition ii, namely:

$$\omega'_W = \omega_B'^{1/p} = (\frac{W'}{B'})^{\frac{p+1}{2p(p-1)}} \quad (30)$$

It is simple to prove that those parameters satisfy the ranges (28), and (27), as well as the initial goal

$$\omega_{B'} < \omega_{W'} \quad (31)$$

Finally, it is possible to find two penalization parameters focused on between class separability ($\omega_W < \omega_B$) by means of considering similar conditions to the previous case, that is:

i. $\omega_W < 1$,|
ii. $\omega_W = \omega_B^p$, p > 1.

These lead to a couple of permissible ranges for ω_B and ω_W given by,

$$0 < \omega_W < 1, \qquad \forall\, p > 1 \quad (32)$$

$$0 < \omega_B < 1 \qquad \forall\, p > 1. \quad (33)$$

It can be noticed that those ranges do not depend on ω_W, ω_B, B′, W′ nor p. Thus, we can choose anyω''_B satisfying the range and theω''_W we chose will be the corresponding value by replacing ω''_B in II, that is $\omega''_B = (\omega''_B)^p$. In order to put an end to the present method, it is advisable to take values between 1−5 for p parameter, as in other cases it may increase the computing time for a SVM training as a result of dealing with high values or it can also produce underfitting due to small values for σ∗ (Wang et al., 2003).

Computational Complexity Analysis

To formally justify the deployment of the proposed analytical criteria in resource-constrained environments, their theoretical computational footprint must be evaluated. The computational complexity of our geometric-differentiable method is strictly governed by the initial pairwise distance extraction required to compute the within-class (W′) and between-class (B′) mean distances. For a training dataset containing N instances and n features, this distance matrix construction operates at a time complexity of $O(N^2 \cdot n)$. Once these static statistical indicators are retrieved, calculating the optimal Gaussian parameter σ^* via the closed-form solutions requires a single algebraic step with a complexity of $O(1)$.

Conversely, traditional hyperparameter tuning paradigms such as exhaustive Grid Search scale exponentially. Grid Search evaluates an optimization matrix across M distinct parameter nodes, requiring the soft-margin SVM to be retrained at every iteration. Since training a standard SVM classifier scales between $O(N^2)$ and $O(N^3)$ based on support vector density, the total time complexity for Grid Search reaches $O(M \cdot N^3)$. Our closed-form variations completely bypass this iterative burden, ensuring real-time operational efficiency.

Results

Before showing some results, we introduce a way to measure the performance of a classifier (SVM) using the evaluation of credit capacity applied to humanitarian logistics. Let $S = \{\overrightarrow{x_1},\overrightarrow{x_2},...,\overrightarrow{x_l}\}$ be a test set, and its corresponding training set $T = \{(\overrightarrow{x_1}, y_1),...,(\overrightarrow{x_l}, y_l)\} \subset R^n \times Y_M$, where YM = {1,2,3,...M}. Its accuracy rate is defined as follows:

$$p_T = \frac{|N_c|}{|N_c| + |N_i|} = \frac{|N_c|}{|S|} \tag{34}$$

where $|\cdot|$ is the set of the points classified correctly $Nc = \{ \vec{x} \in S : (\vec{x}, g(\vec{x})) \in T \}$ and the set of the points classified wrongly $Ni = \vec{x} \in S : (\vec{x}, f(\vec{x})) \in T \}$. Notice that $S = Nc \cup Ni$, where g is the decision function (classifier). The data used in the tests can be found on the repositories of the LIBSVM library web page (Chang & Lin, 2011).

Experimental Setup

To evaluate the practical applicability of the proposed direct- and inverse-variation criteria under restrictive computational conditions, all numerical experiments were intentionally conducted on a low-specification computing environment representative of post-disaster humanitarian logistics scenarios, where centralized computational infrastructure may be unavailable and decision-makers must rely on local resources. The experiments were performed on a computer equipped with a 2.5 GHz processor (4 cores and 4 threads), 4 GB of RAM, and a standard 64-bit operating system, using the LIBSVM framework for implementation. For validation purposes, each dataset was randomly divided into two equally sized subsets, with 50% of the observations used for training and the remaining 50% for testing, allowing the assessment of the proposed parameter-selection criteria under controlled experimental conditions.

Dataset Preparation

Two benchmark datasets available through the LIBSVM repository were considered to evaluate the performance and scalability of the proposed approach: the German Credit Data and ijcnn1 datasets. The former was selected due to its relevance to humanitarian logistics applications, while the latter was used to analyze the computational behavior of the proposed criteria in large-scale classification scenarios. The German Credit Data dataset contains 1000 instances describing socioeconomic and financial characteristics of credit applicants. Since the original dataset includes symbolic attributes, the version employed in this study was transformed into a numerical representation by incorporating indicator variables, enabling its processing by classification algorithms that cannot directly handle categorical information. The dataset contains 20 attributes, including 7 numerical variables associated with credit duration, amount, installment rate, residence duration, age, number of existing credits, and dependents, as well as 13 categorical variables describing factors such as account status, credit history, employment conditions, housing, property ownership, and loan purpose.

The classification objective is to determine whether an applicant belongs to the good-credit class (class 1) or bad-credit class (class 2). Within the humanitarian logistics context, this classification task can be interpreted as a decision-support mechanism for identifying individuals who may require additional assistance during post-disaster reconstruction processes, facilitating a more efficient allocation of financial resources and recovery support.

Experimental Results

First column of the table shows the sort of method used, second column shows the *p* value used, third column contains the ω_B value used for each method, then the optimal σ parameter is shown and the C parameter chosen, as it was mentioned and finally the accuracy rate is shown in the last two columns for training and test sets. Last row shows one of the two sets of penalization parameters proposed by Zhiliang (Liu et al., 2015).

Table 1: LIBSVM, German Credit Data

Method	p	ω_B	σ	C	$^{p}train$	$^{p}test$
$\omega_W = \omega_B^p$	2	0.9	0.2762	2^5	100.0	98.5097
$\omega_W = \omega_B^p$	2	0.8	0.4210	2^5	100.0	98.1258
$\omega_W = \omega_B^p$	4	0.7	0.8987	2^3	100.0	97.912
$\omega_W^p+\omega_B^p=1$	2	ω_B^*	1.8929	2^5	100.0	98.8310
$\omega_W^p+\omega_B^p=1$	5	ω_B^*	1.4021	2^3	100.0	98.4175
Zhiliang	1	0.5	1.699	2^3	100.0	97.982

Table 2: LIBSVM, ijcnn1

Method	p	ω_B	σ	$^{p}train$	$^{p}test$
$\omega_W = \omega_B^p$	2	0.9	0.5556	99.4399	98.7939
$\omega_W = \omega_B^p$	2	0.7	0.3098	99.814	98.1581
$\omega_W = \omega_B^p$	4	0.7	0.1802	99.76	97.7012
$\omega_W^p+\omega_B^p=1$	2	ω_B^*	2.09194	97.695	97.9931
$\omega_W^p+\omega_B^p=1$	6	ω_B^*	2.10039	97.7175	97.99
Zhiliang	1	0.5	2.0899	97.69	97.9793

While Table 1 confirms that the proposed criteria effectively handle classification tracking tasks under standard computational conditions, Table 2 demonstrates their scalability advantage when applied to mass-scale datasets. The ijcnn1 dataset represents a complex large-scale challenge, comprising 141,691 instances, where conventional parameter optimization strategies become computationally demanding. In baseline optimization literature, traditional exhaustive Grid Search for tuning the Gaussian RBF kernel width parameter (σ) requires approximately 6.6 hours of continuous CPU processing. Similarly, intelligent search-based approaches, such as Li's optimization method (Li et al, 2010), significantly reduce this burden but still require around 39 minutes due to their reliance on iterative gradient-based optimization procedures. Such computational delays are impractical in high-pressure scenarios where rapid decision support is essential.

In contrast, the proposed direct- and inverse-variation criteria exploit the structure introduced by Zhiliang's framework (Liu et al., 2015), determining the optimal kernel parameter directly from static class-distance relationships rather than iterative search processes. As a result, the parameter estimation stage is reduced to only 43.4 seconds, representing a computational reduction of approximately 99.8% compared with exhaustive Grid Search. This efficiency is achieved because the proposed criteria preserve the $O(N^2 \cdot n)$ algebraic formulation while introducing a flexible non-linear control exponent (p), overcoming the rigidity of previous closed-form approaches that constrain the penalization parameters to fixed values (e.g., $\omega_W = \omega_B = 0.5$). Nevertheless, although the proposed formulation significantly accelerates the estimation of the kernel parameter (σ), the subsequent optimization of the soft-margin regularization parameter (C) through one-dimensional line search remains comparable between the proposed criteria and Zhiliang's baseline approach. Therefore, the main computational advantage is obtained during the kernel parameter determination stage, where the proposed method eliminates the need for expensive iterative searches.

Overall, Table 2 confirms that this computational acceleration does not compromise predictive performance. Under the direct-variation criterion ($\omega_W = \omega_B^p$) with p=2 and ω_B=0.9, the resulting narrower kernel configuration ($\sigma = 0.5556$) improves local boundary sensitivity and achieves a testing accuracy (p_{test}) of 98.7939%. Likewise, the inverse-variation criterion reaches a peak accuracy of 98.8310%, outperforming Zhiliang's rigid baseline performance of approximately 97.98% and surpassing the maximum accuracy

obtained through conventional Grid Search (98.22%). Therefore, these empirical results demonstrate that the proposed mathematical framework not only preserves the computational efficiency of closed-form approaches but also provides greater adaptability and predictive capability, making it suitable for decision-support systems operating under critical infrastructure and time-constrained humanitarian logistics environment

Discussion and Implications

This research proposes analytical methods for Gaussian Kernel parameter selection in Support Vector Machines (SVM) by generalizing the class separability framework established by Liu et al. (2015). The empirical evaluation using the German Credit dataset indicates that the proposed direct and inverse variation criteria achieve high predictive accuracy, reaching a 98.83% success rate in test sets. By methodically regularizing the penalization parameters (ω_B and ω_W), the proposed approach effectively manages the trade-off between within-class similarity and between-class diversity. This optimization facilitates efficient determination of the hyperparameter σ, avoiding the complexity and high computational time associated with traditional cross-validation or heuristic optimization algorithms (Hsu et al., 2003; Li et al., 2010). These results support the structural risk minimization principle (Vapnik, 1995) by ensuring optimal decision boundaries in high-dimensional kernel spaces (Cover, 1965).

Practical Implications

The practical implications of this research are directly derived from the computational complexity analysis and the execution times obtained during the empirical evaluation. The proposed analytical method achieved a 99.8% reduction in hyperparameter selection time compared with conventional Grid Search, highlighting its potential for deploying decision-support models in environments where computational resources are severely constrained. By eliminating the iterative search bottleneck, the proposed framework enables computationally demanding classification algorithms to operate efficiently on low-specification hardware or isolated local systems. Therefore, the improvement in training efficiency, achieved without compromising predictive accuracy, provides a suitable mathematical alternative for accelerating binary classification processes in time-sensitive decision-making scenarios.

Nevertheless, considering the scope of the experimental validation presented in this study, the implementation of the proposed geometric-differentiable framework in humanitarian logistics and disaster management should be interpreted as a promising future application rather than a fully validated operational solution. Since the experimental assessment was conducted using standardized benchmark datasets, including Statlog (German Credit Data) and ijcnn1, the results demonstrate the methodological robustness of the proposed approach when handling complex and asymmetric data structures; however, they do not directly evaluate its performance under the uncertain and dynamic information flows commonly observed in real disaster environments. Within this context, the German Credit Data dataset can be considered an analogous proxy for evaluating socioeconomic classification problems associated with affected populations. This analytical projection provides a potential pathway for governmental agencies and humanitarian organizations to rapidly assess the socioeconomic conditions and vulnerability levels of affected communities. By reducing computational requirements while maintaining classification performance, the proposed framework could support more efficient allocation of reconstruction resources and assistance programs in future disaster-response applications.

Limitations and Future Research Directions

The study is subject to specific theoretical and methodological limitations. First, the underlying analytical approach adapted from Zhiliang assumes a normal distribution of data to mathematically define the relationship between distances in the input and feature spaces (Liu et al., 2015). Second, the classification performance remains highly sensitive to the exponential tuning parameter (p); an improper selection may distort boundary layouts and lead to underfitting or overfitting, a behavior extensively documented in traditional SVM training paradigms (Wang et al., 2003). Third, while the reported testing accuracy reflects robust global performance, the historical execution script was configured to record only the primary efficiency metric, preventing the retrospective extraction of Precision, Recall, and F1-scores. This exclusive reliance on global accuracy constitutes a significant limitation, particularly given the inherent class imbalance of the German Credit Data dataset and the asymmetric operational costs associated with classification errors in mission-critical domains like humanitarian logistics. In crisis environments, a False Negative carries severe ethical and practical consequences that drastically differ from a False Positive. Consequently, the empirical validation presented in this study must be interpreted strictly as an architectural and computational proof-of-concept focused on algorithmic acceleration, rather than a finalized system optimized for immediate, high-stakes field deployment.

Future research should evaluate the robustness of these analytical methods under highly non-Gaussian distributions to further generalize current structural assumptions. Additionally, future studies must implement exhaustive logging architectures for experimental histories, incorporate cost-sensitive learning frameworks, and employ stratified validation metrics to guarantee safe and stable operations under real-world data imbalances. Finally, subsequent investigations should explore the integration of emerging technologies to improve data traceability in crisis scenarios and extend this geometric-differentiable framework to distance- or instance-based learning algorithms, such as k-nearest neighbor, as well as alternative kernel-based paradigms to further enhance classification performance in resource-constrained environments (Yu et al., 2002).

Conclusions

This study extends Gaussian kernel parameter selection in Support Vector Machines by introducing an adaptive closed-form framework based on direct- and inverse-variation criteria. By incorporating a nonlinear control exponent, the proposed approach overcomes the limitations of previous analytical formulations that rely on fixed penalization configurations. This flexibility enables the model to adjust the balance between within-class compactness and between-class separation according to the characteristics of the dataset, providing a more adaptable mechanism for determining suitable kernel configurations under different classification scenarios.

The empirical evaluation conducted on benchmark datasets demonstrates that the proposed geometric framework significantly reduces the computational cost associated with traditional iterative search procedures while maintaining competitive classification performance. The results confirm its scalability for large-scale classification problems and highlight its potential applicability in decision-support scenarios where computational resources are limited. Although further validation is required under real operational conditions and different data distributions, the findings indicate that closed-form analytical approaches represent a viable alternative for efficient SVM parameter selection, particularly in environments where rapid model configuration and constrained computational infrastructure are relevant considerations.

References

Aronszajn, N. (1950). Theory of reproducing kernels. *Transactions of the American Mathematical Society*, *68*(3), 337–404. https://doi.org/10.2307/ 1990404

Berlinet, A., & Thomas-Agnan, C. (2004). *Reproducing kernel hilbert spaces in probability and statistics*. Springer.

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 27:1–27:27. http://www.csie.ntu.edu.tw/∼cjlin/libsvm

Chapelle, O., Vapnik, V., Bousquet, O., & Mukherjee, S. (2002). Choosing multiple parameters for support vector machines. *Machine Learning*, *46*, 131–159.

Cover, T. M. (1965). Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, *EC-14*(3), 326–334.

Deng, N., Tian, Y., & Zhang, C. (2013). *Support vector machines optimization based theory, algorithms and extensions*. Chapman & Hall/CRC Press.

Goshvarpour, A., & Goshvarpour, A. (2012). Radial basis function and k-nearest neighbor classifiers for studying heart rate signals during meditation. *I.J.Modern Education and Computer Science*, *4*, 43–49.

Griva, I., Nash, S. G., & Sofer, A. (2009). *Linear and nonlinear optimization* (2nd ed.). Society for Industrial; Applied Mathematics.

Herbrich, R. (2002). *Learning kernel classifiers theory and algorithms*. The MIT Press.

Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, *13*(2), 415– 425.

Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). *A practical guide to support vector classification* (tech. rep.). Department of Computer Science, National Taiwan University. Taipei 106, Taiwan. https://www.csie.ntu.edu.tw/ ˜cjlin/

Huang, C. L., Chen, M. C., & Wang, C. J. (2007). Credit scoring with a data mining approach based on support vector machines. *Expert Systems with Applications*, 33(4), 847-856. https://doi.org/10.1016/j.eswa.2006.07.007

Li, C. H., Lin, C. T., Kuo, B. C., & Chu, H. S. (2010). An automatic method for selecting the parameter of the rbf kernel function to support vector machines. *IEEE International conference on Geoscience and Remote Sensing Symposium*, 836–839.

Liu, Z., Zuo, M. J., Zhao, X., & Xu, H. (2015). An analytical approach to fast parameter selection of gaussian rbf kernel for support vector machine. *Information Science and Engineering*, *31*, 691–710.

Mercer, J. (1909). Functions of positive and negative type, and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society of London*, *209*, 415–446.

Schölkopf, B., & Smola, A. J. (2002). *Learning with kernels support vector machines, regularization, optimization, and beyond*. The MIT Press.

Tsoumakas, G., & Katakis, I. (2007). Multi-label classification: An overview. *International Journal of Data Warehousing and Mining*, *3*, 1–13.

Vapnik, V., & Chervonenkis, A. (1964) On a class of algorithms of learning pattern recognition. *Avtomatika i Telemekhanika*, *25*(6).

Vapnik, V. N. (1995). *The nature of statistical learning theory* (2nd ed.). Springer.

Wang, W., Xu, Z., Lu, W., & Zhang, X. (2003). Determination of the spread parameter in the gaussian kernel for classification and regression. *Neurocomputing*, *55*, 643–663.

Widodo, A., & Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing*, 21(6), 2560-2574. https://doi.org/10.1016/j.ymssp.2006.12.007

Yu, K., Ji, L., & Zhang, X. (2002). Kernel nearest-neighbor algorithm. *Neural Processing Letters*, *15*, 147–156.

Zhao, M., Fu, C., Ji, L., Tang, K., & Zhou, M. (2011). Feature selection and parameter optimization for support vector machines: A new approach based on genetic algorithm with feature chromosomes. *Expert Systems with Applications*, *38*, 5197–5204.