STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–30.
Published online in International Academic Press (www.IAPress.org)

Weighted Least Squares Support Vector Machine for Survival Analysis

Rahmawati [1,*], Sri Astuti Thamrin [2,*], Armin Lawi [1], Jeffry Kusuma [1]

[1]*Department of Mathematics, Hasanuddin University, Makassar, Indonesia*
[2]*Department of Statistics, Hasanuddin University, Makassar, Indonesia*

Abstract
Background: The increasing complexity and volume of data across various disciplines have encouraged the use of machine learning methods, including in survival analysis. Given the large percentage of censored data in survival datasets, a methodological technique that can generate more precise survival probability forecasts is required. This study aims to advance survival analysis by applying the Weighted Least Squares Support Vector Machine, using a weighting approach to manage the information imbalance between censored observations and event occurrences. This strategy can yield a prognostic index that is easily categorized into low-risk and high-risk groups.
Methods: This study proposes the Survival Weighted Least Squares Support Vector Machine (Surv-WLSSVM) model through the integration of a weighting strategy based on the Kaplan–Meier estimator. Data with events are assigned weights that consider the value of the survival function, while censored data are given constant weights. Surv-WLSSVM was applied to both simulated and real datasets, and the results were compared with the unweighted method, namely Survival Least Squares Support Vector Machine (Surv-LSSVM), Random Survival Forest (RSF) and Cox Proportional Hazard Regression (CPHR). The simulation scenarios included the complexity of variable numbers, data distribution, sample size, and censoring percentage. The Real datasets used in this study consist of Breastfeeding, PBC, and Bone-Marrow data. A tuning parameters using Particle Swarm Optimization (PSO) was performed to enhance the performance of both Surv-LSSVM and Surv-WLSSVM models. Model performance was evaluated using the concordance index (c-index), where a higher c-index indicates a better model.
Results: In every simulated data setting, the Surv-WLSSVM model continuously showed better performance. Similarly, on real datasets, this model outperformed the alternative and produced more diverse prognostic indices, facilitating the categorization of individuals into low-risk and high-risk groups.
Conclusion: The Surv-WLSSVM represents a significant advancement in SVM-based survival modelling. This approach demonstrates greater reliability and adaptability in handling the complexity of modern survival data. Model performance can be further improved by incorporating a more appropriate weighting scheme for censored observations.

Keywords Prognostic index, Survival Weighted Least Squares Support Vector Machine (Surv-WLSSVM), Kaplan-Meier estimator, Particle Swarm Optimization (PSO), c-index

DOI:10.19139/soic-2310-5070-3016

1. INTRODUCTION

The use of machine learning across various scientific fields has expanded rapidly in line with the increasing complexity and volume of data, including in survival analysis, which requires more accurate and reliable estimation of individual survival probabilities. To date, survival analysis remains one of the primary approaches for addressing problems related to time-to-event outcomes. This approach is widely applied in medicine, industry,

*Correspondence to: Sri Astuti Thamrin (Email: tuti@unhas.ac.id). Department of Statistics, Hasanuddin University, Makassar, Indonesia. Rahmawati (Email: rahmawati.mahuseng@gmail.com).Department of Mathematics, Hasanuddin Unversity, Makassar, Indonesia.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

and economics [1, 2, 3, 4, 5, 6].

In the modern context, the growth in data volume, the diversity of individual characteristics, and high levels of censoring have challenged the performance of conventional methods such as the Cox Proportional Hazards Model. This method has led to the development of machine learning-based approaches, which are more flexible in handling nonlinear relationships, feature interactions, and complex data distributions [2, 7].

Recent studies have demonstrated an increasing trend in the accuracy of survival models through the implementation of machine learning algorithms such as Random Survival Forest (RSF), Artificial Neural Network (ANN), and Support Vector Machine (SVM). RSF has been proven effective in predicting machine failure and the survival of nasopharyngeal cancer [8, 9, 10]. ANN models, including DeepSurv, Cox-nnet, and Nnet-survival, have shown superior performance compared to conventional methods [11, 12, 13, 14].

Meanwhile, SVM has been applied to predict the survival of breast cancer and brain cancer patients by estimating survival functions, hazard rates, and other reliability measures, producing accurate and stable predictions [15, 16]. The development of SVM is recommended to enhance survival analysis due to its advantages over other methods. This method excels in handling data with constant or time-varying event risks [17].

The strength of SVM lies in its ability to handle high-dimensional data and nonlinear patterns [17, 18]. However, most SVM developments have not explicitly accounted for heterogeneous risk structures among individuals, particularly between group censored and event data. Generally, survival data are imbalanced because the proportion of censored data is typically higher than that of event data. This type of imbalance is observed across multiple contexts, including cancer [3, 9, 19, 20, 21], chronic diseases [22, 23], and business failure [4, 24]. This condition may cause models to be overtrained on censored data, and thereby reducing their sensitivity to the event group.

Several developments of Survival SVM have been proposed, one of which is the Survival Least Squares Support Vector Machine (Surv-LSSVM) [25, 26, 27]. However, most existing models have not optimally addressed the differences in contributions between censored and event data. Weighted SVM approaches for survival analysis have been suggested [17], but no research has yet integrated them into the Surv-LSSVM framework, particularly for high-dimensional data with varying censoring levels. Furthermore, the types of weighting used in previous studies remain limited. To date, no studies have evaluated the use of the Kaplan-Meier estimator as a weighting function in Survival SVM. This research gap motivates the development of the Survival Weighted Least Squares Support Vector Machine (Surv-WLSSVM), which integrates Surv-LSSVM with Kaplan-Meier-based weighting.

The novelty of this study lies in the integration of weights derived directly from data through the Kaplan-Meier estimator, enabling individuals with higher event risk to contribute more significantly during model training. This approach makes the model more adaptive to imbalanced risk structures, improves prediction accuracy, and ensures robustness to varying censoring levels. Model performance evaluation was conducted by comparing Surv-LSSVM and Surv-WLSSVM on both simulated and real datasets, with simulation scenarios that considered variations in data distribution, data dimension, and censoring percentage. Model quality was assessed using the concordance index (c-index), where higher values indicate better predictive performance.

2. LITERATURE REVIEW

2.1. Kaplan-Meier Survival Analysis

Several statistical approaches have been used to optimize the performance of survival analysis, including the nonparametric Kaplan-Meier approach, the semiparametric Cox Proportional Hazards Regression, and the Parametric Linear Regression. Approaches to improve the predictive accuracy of survival analysis have also evolved with the modification of several machine learning methods, such as gradient boosting and support vector machines.

The Kaplan-Meier estimator, as a nonparametric approach, remains one of the most widely used methods, particularly in health-related applications [10, 21, 28]. The characteristics of the Kaplan-Meier estimator are as follows: (a) time intervals are often unequal, and calculations are performed at each failure, and (b) the survival function for a given time interval is the proportion of the number of surviving subjects at the beginning of the interval minus the number of failures within that interval. Mathematically, the Kaplan-Meier estimator is given by Equation (1):

$$\hat{S}(t) = \prod_{t_i \leq t} \left(\frac{N_{i-1} - d_i}{N_{i-1}} \right) \tag{1}$$

where
$\hat{S}(t)$: probability that an individual survives longer than time t
t_i : time at which the event occurs
N_{i-1} : number of individuals at risk at time t_i
d_i : number of events at time t_i

The number of censorings or withdrawals is not included in the Kaplan-Meier estimation. The Kaplan-Meier estimator is commonly used to directly estimate the survival function, verify events within each group over a specific period, and test differences between groups [21, 29].

2.2. Cox Proportional Hazard Regression

Survival data modeling using the Cox Proportional Hazard Regression (CPHR) model is a semi-parametric approach that can be employed to describe the relationship between a survival time response variable and one or more predictor variables. The general form of the Cox Proportional Hazard model is expressed as follows:

$$h(t|X) = h_0(t) \exp(\beta X) \tag{2}$$

For covariate variables (X), the model can be written as:

$$h_i(t) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p) \tag{3}$$

where:

- $h_i(t)$: hazard function for the i-th individual
- $h_0(t)$: baseline hazard

- $\beta_1, \beta_2, \ldots, \beta_p$: regression coefficient
- $x_1, x_2, \ldots, x_p$: covariates

The measure used to quantify the relative risk between individuals is known as the *Hazard Ratio (HR)*.

The Cox proportional hazards model is the standard technique widely applied to analyze data with failure-time endpoints by incorporating multiplicative interactions between binary variables. However, in small-sample studies, the use of such standard statistical techniques may lead to bias due to limited sample size or low statistical power, particularly in biomarker data [30, 31]. To date, the CPHR model remains one of the most commonly used methods [32, 33] and holds strong potential for further development through machine learning-based approaches [33, 34, 35].

2.3. Random Survival Forest

The Random Survival Forest (RSF) is a modification of the traditional Random Forest (RF) algorithm specifically designed to handle survival data. This method employs a machine learning approach based on decision trees to analyze right-censored time-to-event data [36]. An ensemble of decision trees works collectively to model survival data [37].

The implementation of the RSF algorithm consists of the following steps [38]:

1. Bootstrap sampling: Draw (S) bootstrap samples from the original dataset with replacement. Each bootstrap sample is used to construct a survival tree. Typically, each bootstrap sample includes approximately 63% of the original data, leaving out around 37%, referred to as Out-of-Bag (OOB) data.
2. Tree construction: Grow a survival tree for each bootstrap sample. At each node, randomly select (p) candidate variables. The node is split using the variable that maximizes the survival difference between the child nodes, incorporating survival time and censoring information.
3. Full tree growth: Grow each tree to its full size, with the constraint that each terminal node must contain at least $(d_0 > 0)$ unique failure events.
4. Cumulative Hazard Function (CHF): Calculate the CHF for each tree, and then compute the ensemble CHF by averaging the CHFs across all trees in the forest.
5. Prediction error estimation: Compute the prediction error of the ensemble CHF using the OOB data. This prediction error is then used to determine variable importance.

Applications of the RSF algorithm have demonstrated superior performance in several areas, including survival rate prediction and risk stratification of patients with gastric neuroendocrine neoplasms [39], early prediction of Alzheimer's disease [40], and survival analysis in cancer, diabetes, and kidney transplantation studies. These studies have shown that RSF outperforms the Cox Proportional Hazard Regression (CPHR) model, providing accurate, quantifiable, and interpretable risk estimates [32, 33, 36, 41].

2.4. SVM: Support Vector Machine

Support Vector Machine (SVM) was originally introduced as a classification method. Its high level of accuracy has led to its application in many fields [42, 43, 44, 45]. The basic concept of SVM is solving an optimization problem

that involves an objective function and constraints. Mathematically, for a dataset $\{(\mathbf{x}_i, y_i)\}_{i=1}^n$, with $\mathbf{x}_i \in \mathbb{R}^d$ and $y_i \in \{-1, 1\}$, SVM solves the optimization problem with the objective function defined in Equation (4):

$$\min_{w,b,\xi} \ \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{n}\xi_i \tag{4}$$

subject to

$y_i(\mathbf{w}^T\boldsymbol{\varphi}(\mathbf{x}_i) + b) \geq 1 - \xi_i,\ \xi_i \geq 0, i = 1, 2, \ldots, n.$

where $\boldsymbol{\varphi}(\cdot)$ is the mapping function into a high-dimensional feature space, C is the regularization parameter, and ξ_i is the slack variable.

The SVM solution is obtained by solving a system of linear equations, replacing quadratic programming through the application of kernel functions. Several kernel functions can be used, including the linear kernel, Radial Basis Function (RBF), and polynomial kernels.

2.5. LS-SVM: Least Squares Support Vector Machine

The SVM method is a reliable and accurate approach and is capable of handling degradation disturbances in classification problems [46]. One limitation of SVM is its high computational complexity due to quadratic programming and its sensitivity to outliers, since only support vectors influence classification decisions [47]. As a solution, the Least Squares Support Vector Machine (LS-SVM) was developed by modifying the SVM objective function and constraints into a least-squares form. LS-SVM replaces the loss function with a quadratic form and the constraints with equality expressions. The LS-SVM objective function is shown in Equation (5).

$$\min_{w,b,\xi} \ \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}C\sum_{i=1}^{n}\xi_i^2 \tag{5}$$

subject to $y_i(\mathbf{w}^T\boldsymbol{\varphi}(\mathbf{x}_i) + b) = 1 - \xi_i; i = 1, 2, \ldots, n.$

By applying Lagrangian multipliers and eliminating variables, the LS-SVM solution can be obtained by solving a system of linear equations rather than quadratic programming, which is computationally more efficient [44, 48], and also demonstrates improved performance [49].

2.6. WLS-SVM: Weighted Least Squares Support Vector Machine

The LS-SVM model assigns equal weights to all data points, making it sensitive to outliers or non-representative data. To address this limitation, the Weighted Least Squares Support Vector Machine (WLS-SVM) was introduced. This method incorporates a weight B_i for each data point, reflecting its relative importance in the learning process. The mathematical formulation of WLS-SVM is:

$$\min_{w,b,\xi} \ \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}C\sum_{i=1}^{n}B_i\xi_i^2 \tag{6}$$

subject to $y_i(\boldsymbol{\varphi}(\mathbf{x}_i)^T\mathbf{w} + b) = 1 - \xi_i$.
The weights B_i are often computed using robust techniques such as fuzzy clustering or kernel-based possibilistic

c-means, such that outliers receive small weights and do not dominate the solution [19, 47].

The WLSSVM approach not only preserves the computational efficiency of LS-SVM but also improves the model's ability to handle noisy or outlier-contaminated data. Recent studies have shown that WLS-SVM excels in describing complex data distribution trends, particularly when data exhibit heterogeneous distributional directions [19, 50].
The application of both LS-SVM and WLS-SVM methods demonstrates that the models can be easily implemented, provide excellent generalisation results, and incur relatively low computational costs [44, 45, 50, 51, 52, 53]. LS-SVM has been proven to be the most consistent in terms of accuracy, stability, and responsiveness to preprocessing techniques [48]. However, in practice, the sparsity property is lost in LS-SVM, and support value estimation is only optimal when the error variable follows a Gaussian distribution. Both of these issues can be addressed through the weighting strategy in the WLSSVM method [54].

2.7. Kernel Functions in SVM

Kernel functions are a fundamental concept in SVM methods. A kernel $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ can be interpreted as an inner product in a high-dimensional feature space, expressed as $K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle,\ x_i, x_j \in \mathcal{X}$. Commonly used kernel functions include polynomial, sigmoid, and Radial Basis Function (RBF) kernels [55]:

Table 1. Types of Kernel Functions

Kernel Function	Functional Form
Polynomial	$K(x_i, x_j) = (x_i^T x_j + \theta)^p$ where p is a positive integer, $\theta \in \mathbb{R}$.
Sigmoid	$K(x_i, x_j) = \tanh(\alpha x_i^T x_j + \theta)$ where $\alpha \in \mathbb{R}$, and $\theta \in \mathbb{R}$.
RBF	$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$ where $\sigma > 0$ is the kernel parameter.

The performance of SVM is highly dependent on the selection of two main parameters, namely the regularization parameter C and the kernel parameter σ, particularly when using the RBF kernel function. Suboptimal parameter selection can degrade the model's generalization ability [56]. Parameter tuning in SVM is generally performed using the grid search method. Alternative approaches that are more computationally efficient can also be applied to ensure better solutions.

2.8. Surv-SVM: Survival Support Vector Machine

The SVM method for survival analysis is implemented on high-dimensional survival data. The basic concept of Surv-SVM analysis is to determine a prognostic function. In the context of the Cox model, is this function serves as a linear predictor that forms a mathematical framework for modelling the relative risk (hazard ratio) based on covariates. When this function is evaluated on the data of a specific individual, it yields a numerical value called

the prognostic index. This index reflects the individual's relative log-hazard, providing a risk score that can be used for risk group stratification and further analysis [57].

The prognostic function is generally used to predict the future outcome or progression of a condition based on currently available data or information. This term is often used in the medical field to estimate the potential outcomes of a disease [10, 12, 58]. Mathematically, the prognostic function in the Surv-SVM model is defined as a function $u : \mathbb{R}^d \to \mathbb{R}$, which in matrix form is given in Equation (7):

$$\mathbf{u}(\mathbf{x}) = \mathbf{w}^T \boldsymbol{\varphi}(\mathbf{x}) \tag{7}$$

where $\mathbf{w}$ is the parameter vector and $\varphi(\mathbf{x})$ is a transformation of the covariate variables $\mathbf{x}$ [25, 26, 27, 59].

Given a dataset $D = \{(\mathbf{x}_i, t_i, \delta_i)\}$ with $\mathbf{x}_i$, t_i, and δ_i representing the predictor variables, survival time, and censoring status of the i-th subject, respectively, the support vector machine model for survival analysis is expressed as follows:

$$\min_{\mathbf{w},\xi} \quad \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}C\sum_i \sum_{i,j} r_{ij}\xi_{ij}; C \geq 0 \tag{8}$$

subject to the constraint $\mathbf{w}^T\varphi(\mathbf{x}_j) - \mathbf{w}^T\varphi(\mathbf{x}_i) \geq 1 - \xi_{ij}; \forall i < j$
where r_{ij} is defined as:

$$r_{ij} = \begin{cases} 1; & (t_i < t_j, \delta_i = 1) \\ 0; & (t_i \geq t_j, \delta_i = 0) \end{cases}$$

which serves as the comparison indicator between the i-th object $(\mathbf{x}_i, t_i, \delta_i)$ and the j-th object $(\mathbf{x}_j, t_j, \delta_j)$ [25, 26].

2.9. Surv-LSSVM: Survival Least Square Support Vector Machine

The Surv-LSSVM model is an extension of survival analysis using the LS-SVM method. The objective function and the constraint function of the Surv-LSSVM model are as follows:

$$\min_{w,\xi} \quad \frac{1}{2}\mathbf{w}^\top\mathbf{w} + \frac{1}{2}C\sum_{i=1}^{n}\sum_{i<j}^{n} r_{ij}\,\xi_{ij}^2 \tag{9}$$

with the corresponding constraint $\mathbf{w}^T\varphi(\mathbf{x}_j) - \mathbf{w}^T\varphi(\mathbf{x}_i) \geq 1 - \xi_{ij}; \forall i < j$.

The solution to this equation is obtained by using a linear approach in solving the KKT (Karush–Kuhn–Tucker) system. This system is a generalization of the Lagrange method for solving optimization problems with equality and inequality constraints [27, 59].

The prognostic function of the Surv-LSSVM model is shown in Equation (10):

$$\mathbf{u} = \boldsymbol{\alpha}^T\mathbf{D}\mathbf{K}_n(\mathbf{x}) \tag{10}$$

where $\boldsymbol{\alpha} = (C\mathbf{DKD}^T + \mathbf{I})^{-1}C\mathbf{1}$, C is the regularization parameter, matrix $\mathbf{D}$ has elements {-1,0,1}, which are defined such that they satisfy the following equation:

$$\mathbf{DX} = \begin{bmatrix} x_1 - x_2 \\ \vdots \\ x_1 - x_n \\ \vdots \\ x_{n-1} - x_n \end{bmatrix} \tag{11}$$

with $\mathbf{X} = (x_1, x_2, \ldots, x_n)^T$ and $\mathbf{D}$ of dimension $m \times n$, where m is the number of comparable objects and n is the number of observed objects [59, 25].

The matrix $\mathbf{K}$ is an $n \times n$ kernel matrix, each element of which contains the kernel function value between individual i and j, denoted as $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2)$.

I is an $n \times n$ identity matrix, and $\mathbf{1}$ is an $n \times 1$ column matrix with all elements equal to 1.

2.10. Tuning Parameters using Particle Swarm Optimization

Recent studies have proposed parameter tuning processes using a metaheuristic approach based on the Particle Swarm Optimization (PSO) algorithm. This method is a population-based algorithm that mimics the behavior of a swarm in finding the optimal location within the search space. In the context of SVM parameter tuning, PSO is employed to find the pair (C, σ) that minimizes the error function of the SVM model [60].

The use of the Particle Swarm Optimization algorithm to tune parameters in the SVM model with an RBF kernel has been shown to produce more accurate and efficient predictive models compared to SVM without automatic tuning [45]. This approach avoids bias due to fixed parameters and enhances the model's ability to adapt to the complexity of the data.

PSO has gained popularity in various studies due to its fast convergence speed, low computational cost, and flexibility in handling different types of objective functions in both statistical and machine learning models [55]. Moreover, PSO is capable of solving nonlinear optimization and complex parameter adjustment problems, consistently providing more stable and superior solutions compared to classical numerical methods or genetic algorithms [61, 62, 63].

The basic concept of the PSO algorithm is inspired by the social behavior of animals such as bird flocks. The search begins with a random population called particles, each having a velocity that is continuously updated. The velocity of each particle is based on both individual and collective best experiences, gradually moving toward better solutions [62].

The steps of the PSO algorithm are described as follows:

1. Initialize particles and their velocities in a swarm.
2. Evaluate particles by comparing fitness values, which are used to determine the Global best (G_b) and Personal best (P_b).
3. Update P_b by comparing its value before and after an iteration; if the new particle's fitness is higher than the previous P_b, then it becomes the new P_b.

4. Determine G_b as the best particle among all members of the swarm, i.e., the particle with the highest fitness value.
5. Update the velocity of each particle:

$$v_i^{t+1} = \theta v_i^t + \lambda\tau_1[P_b - z_i^t] + \beta\tau_2[G_b - z_i^t] \tag{12}$$

v_i^{t+1}	: new velocity
v_i^t	: initial velocity, defined as 0 at $t = 0$
θ	: inertia weight, constant, typically $0.5 \leq \theta \leq 0.9$
τ_1 and τ_2	: random numbers in the interval $[0, 1]$
λ and β	: acceleration coefficients, commonly $\lambda \approx \beta \approx 2$
P_b	: Personal best
G_b	: Global best

6. Update the particle position using $z_i^{t+1} = z_i^t + v_i^{t+1}$.
7. Repeat step 2 until the fitness reaches a sufficiently good value or the predefined maximum iteration is met.

The inertia weight controls the influence of the previous velocity on the current velocity and plays an important role in PSO as it regulates the balance between local and global exploration [63]. As iterations proceed, the swarm model is strategically guided to focus the search on high-quality regions of the solution space [61]. Each particle moves within the solution space with the ability to recall and preserve its best previous position across iterations. This working mechanism makes the PSO algorithm more adaptive and efficient.

2.11. Concordance Index

The concordance index, commonly denoted as the c-index, is an evaluation metric used to measure the predictive ability of models in time-to-event data, particularly in survival analysis. The c-index assesses how well a model can predict the correct ordering of event times, while taking censored data into account.

Survival data often contain censored individuals, meaning that the event time is unknown for some cases. As a result, pairs involving censored individuals cannot be compared if their event order cannot be determined. Therefore, the c-index is designed to address this issue by considering only comparable pairs [25, 26].

$$\textit{c-index} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} r_{ij} I(u(x_j) > u(x_i))(t_j - t_i) > 0)}{\sum_{i=1}^{n}\sum_{j=1}^{n} v_{ij}} \tag{13}$$

where I is the indicator function, $u(\mathbf{x})$ is the prognostic index, and t is the survival time.

2.12. State-of-the-Arts

Several previous studies have shown that survival analysis has undergone significant model development. Machine learning approaches have become popular and proven to improve the performance of survival analysis. This study proposes the development of survival analysis using the Weighted Least Square Support Vector Machine. Related studies are summarized as follows.

1. Studies on the application of SURLS-SVM demonstrated that this method can overcome the limitations of the Cox PHM, as it does not require the proportional hazards assumption, and it produces a higher c-index in both simulated and clinical data [27].

2. Research on A-SURLSSVM provided empirical evidence that this new LSSVM variant consistently outperformed the Cox model in both simulated and clinical datasets, although its performance decreased when the censoring percentage was high [26].

3. Further studies on the development of SVM for survival data provided a theoretical basis for the suitability of SVM for high-dimensional data and recommended the use of weighting strategies for censored observations [17].

4. Research on principal WLSSVM showed that WLSSVM is flexible and can be extended to various domains, including survival analysis, thereby strengthening the foundation for the development of WLSSVM-based models [50].

5. Papers on WLSSVM provided an important methodological basis for Surv-WLSSVM, demonstrating its ability to capture complex data structures, computational efficiency, and the potential of using weights to address censoring [19].

6. Studies on censored patients in Kaplan–Meier analysis revealed bias due to informative censoring, supporting the need to integrate weights into survival models to produce more accurate estimates [64].

Figure 1 presents the State-of-the-Art diagram of the research, which illustrates the foundation for the development of the Surv-WLSSVM method. Based on this review, it can be concluded that the development of Support

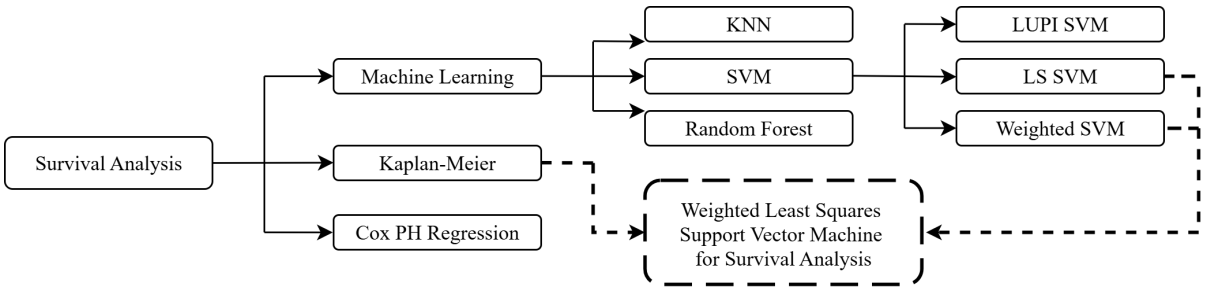


Figure 1. State-of-the-Art Diagram of the Study

Vector Machine-based methods, particularly Surv-WLSSVM, shows strong potential for superior performance. This method not only addresses limitations of assumptions and handles high-dimensional data, but also provides more accurate survival estimates across various scenarios. Nevertheless, further exploration remains necessary, especially in the selection of parameters and handling datasets with varying levels of censoring. Therefore, further research on Surv-WLSSVM is essential to broaden its application to both clinical and more complex simulated data.

3. Research Methods

3.1. Data Sources

This study utilizes two types of data sources, namely simulated data and real-world data. Simulated data are generated to replicate the characteristics of real data, particularly in terms of the number of predictor variables. The use of simulated data aims to systematically evaluate the performance of the model. Through simulation, various scenarios are constructed with variations in the number of variables, sample sizes, and censoring rates to assess the

consistency of model performance. The selection of the number of variables and sample size in the simulated data is based on several key considerations. The characteristics and number of variables are designed to mimic real-world data conditions, ensuring that the analysis results are relevant and representative of actual situations. The sample size (n) is varied to reflect studies with small, medium, and large datasets, allowing model performance to be evaluated across different data scales. Additionally, the percentage of censored cases is determined based on empirical data to assess model performance under low, moderate, and high censoring levels. This approach ensures that the simulated data accurately capture the variability and complexity typically encountered in survival studies.

Table 2. Distribution of simulated data variables

Dataset	Variables	Distribution	Parameters
12 variables	X_1, X_2, X_3, X_4	$N(\mu, \sigma^2)$	$\mu_1 = 45, \mu_2 = 50, \mu_3 = 55, \mu_4 = 60,$ $\sigma_1^2 = 2, \sigma_2^2 = 4, \sigma_3^2 = 6, \sigma_4^2 = 8$
	X_5, X_6, X_7, X_8, X_9	Poisson(λ)	$\lambda_5 = 2, \lambda_6 = 3, \lambda_7 = 4, \lambda_8 = 5, \lambda_9 = 6$
	X_{10}, X_{11}, X_{12}	Binomial(s, p)	$s_{10} = s_{11} = s_{12} = 1, p_{10} = 0.2, p_{11} =$ $0.3, p_{12} = 0.4$
17 variables	X_1, X_2, X_3, X_4	$N(\mu, \sigma^2)$	$\mu_1 = 45, \mu_2 = 50, \mu_3 = 55, \mu_4 = 60,$ $\sigma_1^2 = 2, \sigma_2^2 = 4, \sigma_3^2 = 6, \sigma_4^2 = 8$
	X_5, X_6, X_7, X_8, X_9	Poisson(λ)	$\lambda_5 = 2, \lambda_6 = 3, \lambda_7 = 4, \lambda_8 = 5, \lambda_9 = 6$
	$X_{10}, X_{11}, X_{12}, X_{13}$	Binomial(s, p)	$s_{10} = s_{11} = s_{12} = s_{13} = 1, p_{10} =$ $0.2, p_{11} = 0.3, p_{12} = 0.4, p_{13} = 0.5$
	$X_{14}, X_{15}, X_{16}, X_{17}$	$\Gamma(\alpha, \beta)$	$\alpha_{14} = \alpha_{15} = \alpha_{16} = \alpha_{17} = 2,$ $\beta_{14} =$ $0.2, \beta_{15} = 0.4, \beta_{16} = 0.6, \beta_{17} = 0.8$
35 variables	$X_1, X_2, X_3, X_4, X_5, X_6, X_7$	$N(\mu, \sigma^2)$	$\mu_1 = 45, \mu_2 = 50, \mu_3 = 55, \mu_4 = 60, \mu_5 =$ $65, \mu_6 = 70, \mu_7 = 75,$ $\sigma_1 = \sigma_2 = 2, \sigma_3 =$ $\sigma_4 = 4, \sigma_5 = \sigma_6 = 6, \sigma_7 = 8$
	$X_8, X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{14}$	Poisson(λ)	$\lambda_8 = 2, \lambda_9 = 3, \lambda_{10} = 4, \lambda_{11} = 5, \lambda_{12} =$ $6, \lambda_{13} = 7, \lambda_{14} = 8$
	$X_{15}, X_{16}, X_{17}, X_{18}, X_{19}, X_{20}, X_{21}$	Log Normal(μ_l, σ^2)	$\mu_{15} = 0, \mu_{16} = 0.1, \mu_{17} = 0.2, \mu_{18} =$ $0.3, \mu_{19} = 0.4, \mu_{20} = 0.5, \mu_{21} = 0.6,$ $\sigma_{15} =$ $\sigma_{16} = \sigma_{17} = \sigma_{18} = \sigma_{19} = \sigma_{20} = \sigma_{21} = 1$
	$X_{22}, X_{23}, X_{24}, X_{25}, X_{26}, X_{27}, X_{28}$	Binomial(s, p)	$s_{22} = s_{23} = s_{24} = s_{25} = s_{26} = s_{27} =$ $s_{28} = 1,$ $p_{22} = 0.2, p_{23} = 0.3, p_{24} =$ $0.4, p_{25} = 0.5, p_{26} = 0.6, p_{27} = 0.7, p_{28} =$ 0.8
	$X_{29}, X_{30}, X_{31}, X_{32}, X_{33}, X_{34}, X_{35}$	$\Gamma(\alpha, \beta)$	$\alpha_{29} = \alpha_{30} = \alpha_{31} = \alpha_{32} = \alpha_{33} = \alpha_{34} =$ $\alpha_{35} = 2,$ $\beta_{29} = 0.1, \beta_{30} = 0.2, \beta_{31} =$ $0.3, \beta_{32} = 0.4, \beta_{33} = 0.5, \beta_{34} = 0.6, \beta_{35} =$ 0.7

.

In this study, the performance of the Surv-WLSSVM model with weighting is compared to the Surv-LSSVM model without weighting. The performance of both models is measured using the c-index, where a larger c-index indicates better model performance. This approach provides a strong foundation before applying the model to more complex real datasets. The number of variables in the simulated data mirrors the number of variables in the real data used in this study. Simulated data are generated from specific distributions. Variations in the number of variables consist of 12, 17, and 35 variables, with sample sizes of 100, 300, 500, and 1000, respectively. Furthermore, variations in the proportion of censored data range from 10%, 20%, up to 90% of the total data. In total, 108 simulated datasets were generated for this study.

The simulated data in this study employ survival times generated from a Weibull distribution, while censoring status is randomly generated as 0 or 1. The predictor variables are generated from specific distributions as described in Table 2.

The Surv-LSSVM and Surv-WLSSVM models are applied to simulated data by randomly generating several combinations of censored-to-event ratios, ranging from 10% : 90%, 20% : 80% up to 90% : 10%. Model performance for each simulated dataset is evaluated using the c-index.

Furthermore, the real-world datasets used in this study comprise Breastfeeding, PBC, and Bone-Marrow data, each containing 12, 17, and 35 predictor variables, respectively. The characteristics of these predictor variables are constructed using distributions such as the normal, Poisson, binomial, log-normal, and gamma. The three real-world datasets are described as follows:

1. **Breastfeeding Data**: consists of 12 variables and 718 individuals. This dataset records breastfeeding practices by mothers since childbirth, obtained from the Indonesian Family Life Survey wave 5 (IFLS-5) and available at https://www.rand.org. After preprocessing, the dataset contains 718 individuals. It includes both response and predictor variables. The response variable represents the duration of exclusive breastfeeding since birth. Mothers who introduced water, other foods, or weaned their infants before six months are considered as experiencing an event, while mothers who exclusively breastfed for a full six months are categorized as censored data [65].

2. **PBC**: The Primary Biliary Cholangitis (PBC) dataset comprises 17 predictor variables and 258 individuals. PBC is a progressive chronic liver disease primarily affecting the small intrahepatic bile ducts. Progressive damage leads to impaired bile flow, inflammation, and eventually cirrhosis. In advanced stages, the disease results in liver failure and death if untreated, with liver transplantation often being the only life-saving therapy [66]. In this dataset, patients who died during the study are classified as events, while patients who survived without transplantation until the end of the study are considered censored.

3. **Bone Marrow**: consists of 35 variables and 182 individuals. The event status corresponds to patients who died during the study, while the censored status corresponds to patients who survived until the end of the study. Bone marrow is a soft tissue located within the cavities of bones and serves as the primary site of blood cell production. With ageing, changes in blood cell composition and immune responses occur, increasing susceptibility to age-related hematologic conditions such as anaemia, immunodeficiency, and hematologic malignancies [67].

The evaluation of Surv-WLSSVM performance on both simulated and real-world datasets aims to demonstrate the model's robustness under varying levels of data complexity. Model performance is measured using the c-index, where higher values indicate better performance.

3.2. Research Workflow

This study was conducted through structured stages to achieve the research objectives effectively. The steps include problem identification and formulation of the weighted Surv-LSSVM development, referred to as Surv-WLSSVM. The model was implemented on both simulated and real-world datasets. Implementation involved parameter tuning using the PSO algorithm and calculating the c-index. A comparison of the performance between Surv-WLSSVM, Surv-LSSVM, RSF and CPHR was then carried out based on c-index evaluation. The final stage involved drawing conclusions and discussing directions for future research.

The research workflow is outlined as follows:

1. Generation of simulated data consisting of three categories with 12, 17, and 35 predictor variables, respectively. The sample sizes considered are 100, 300, 500, and 1000. The censoring proportions range from 10%, 20%, up to 90%, resulting in a total of 108 datasets. This design aims to mimic realistic survival data with diverse characteristics.
2. Implementation of the model on simulated data to evaluate the performance of the Surv-WLSSVM model using Kaplan–Meier based weighting for event observations and several constant weighting schemes for censored observations. The model implementation consists of the following steps:
   (i) Solving the mathematical formulation of the Surv-WLSSVM model.
   (ii) Computing weights for event observations using the Kaplan–Meier estimator.
   (iii) Determining constant weights for censored observations through simulations using several values, namely 0.1, 0.25, 0.5, and 0.75.
   (iv) Tuning the parameters of the Surv-WLSSVM and Surv-LSSVM models using the PSO algorithm.
   (v) Evaluating the c-index; the weight pair in step (iii) that yields the highest c-index is selected as the weighting scheme for the model.
3. Comparison of the Surv-WLSSVM, Surv-LSSVM, RSF, and CPHR models to assess the relative performance of the weighted Surv-WLSSVM against the unweighted Surv-LSSVM and two commonly used survival analysis methods, RSF and CPHR.
4. Application of the models to real data, including parameter tuning using the PSO algorithm. The optimal parameters are determined, and the trends of the parameters and resulting c-index values are described.
5. Comparison of Surv-WLSSVM, Surv-LSSVM, RSF, and CPHR models on real data to identify the best-performing model based on the highest observed c-index.
6. Drawing conclusions and interpreting the performance of the Surv-WLSSVM model constructed using the Kaplan–Meier–based weighting estimator.

In general, the research workflow is presented in Figure 2.

Parameter tuning in Surv-WLSSVM is a fundamental step that significantly influences prediction accuracy. The use of the PSO algorithm aims to select the optimal regularization parameter C and kernel parameter σ, i.e., the parameter values that maximize the c-index. In the PSO algorithm, each particle in the swarm represents a candidate solution in the form of a parameter pair (C, σ).

The specific steps of the PSO algorithm for tuning parameters C and σ in Surv-WLSSVM are described as follows:

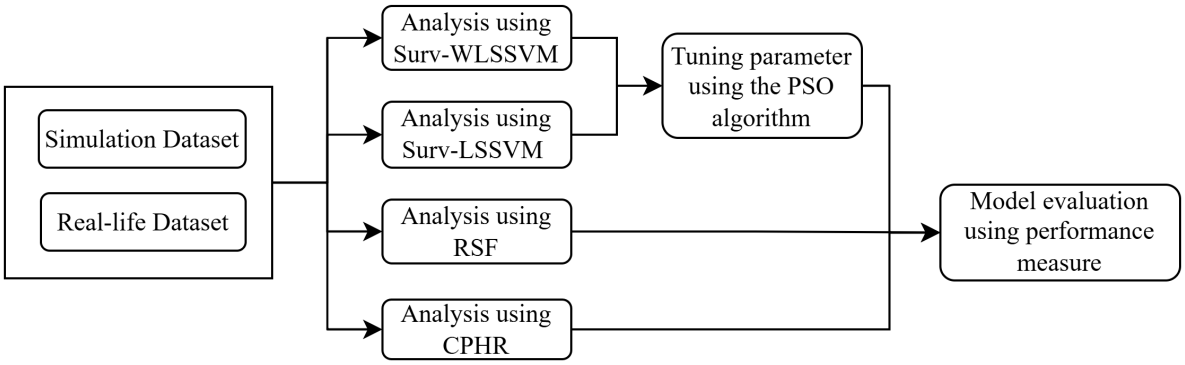


Figure 2. Research Flowchart

1. Input survival data $D = \{(\mathbf{x}_i, t_i, \delta_i)\}$
2. Initialize the initial position of particles $z_i = (C_i, \sigma_i)$ with velocity $v_i = (v_{C_i}, v_{\sigma_i})$
3. Identify $P_{bi} = (C_i^*, \sigma_i^*)$ and $G_b = (C^*, \sigma^*)$
4. Update particle velocity and position using the updating equation:
   Particle velocity:

$$v_c^{t+1} = \theta v_c^t + \lambda\tau_1[C_i^* - C_i^t] + \beta\tau_2[C^* - C_i^t] \tag{14}$$
$$v_\sigma^{t+1} = \theta v_\sigma^t + \lambda\tau_1[\sigma_i^* - \sigma_i^t] + \beta\tau_2[\sigma^* - \sigma_i^t] \tag{15}$$

   where the inertia weight θ is set within the range of $\theta \approx 0.5 \sim 0.9$, τ_1 and τ_2 are random numbers drawn from the interval $[0, 1]$, and the acceleration coefficients λ and β are typically assigned $\lambda \approx \beta \approx 2$.

   Particle position:

   The position of each particle is updated using the following equations:

$$C_i^{t+1} = C_i^t + v_c^{t+1} \tag{16}$$
$$\sigma_i^{t+1} = \sigma_i^t + v_\sigma^{t+1} \tag{17}$$

   where C_i^t and σ_i^t represent the position parameters of the i-th particle at iteration t, and v_c^{t+1} and v_σ^{t+1} denote the updated velocity components.

5. Evaluate fitness with the following steps:
   (i) Compute the matrix $\mathbf{D}_{(m\times n)}$
   (ii) Compute the kernel matrix $\mathbf{K}_{(n\times n)}$ using the RBF kernel function and kernel parameter σ_i^{t+1} obtained in step 4
   (iii) Compute the weight matrix using the Kaplan–Meier estimator as shown in Equation (1)
   (iv) Compute the matrix $\boldsymbol{\alpha}$

(v) Compute the prognostic index $\mathbf{u}$

(vi) Compute the c-index using Equation (13)

6. Repeat iterations from steps 2–5, continuously updating particle positions based on c-index evaluation results. Iterations stop when the maximum number of iterations is reached, or when there is no significant improvement in c-index.

The PSO algorithm efficiently optimizes the parameters of Surv-WLSSVM through adaptive exploration and exploitation in the parameter space. Evaluation based on the c-index ensures that the resulting parameters not only minimize prediction errors but also improve the predictive ranking accuracy of the survival model.

4. RESULTS

4.1. The Resulted Models

WLS-SVM is an extension of traditional SVM that introduces weights to the training data in order to assign higher priority to certain data considered more important or with higher confidence values. The basic idea of WLS-SVM is to minimise a loss function that accounts for the varying weights of the data. The objective function in WLS-SVM consists of two main components: (i) the error term, which measures the discrepancy between the model prediction and the actual values, and (ii) the regularisation term, which minimises model complexity. The objective function and constraints of the SURV-WLSSVM model are given in Equation (18):

$$\min_{w,\xi} \quad \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}C\sum_{i=1}^{n}\sum_{j=1}^{n} r_{ij}B_{ij}\xi_{ij}^2 \tag{18}$$

subject to

$\mathbf{w}^T\boldsymbol{\varphi}(\mathbf{x}_j) - \mathbf{w}^T\boldsymbol{\varphi}(\mathbf{x}_i) = 1 - \xi_{ij}, \forall i, j = 1, 2, \ldots, n.$

B_{ij} represents the weight component for the i-th object $(\mathbf{x}_i, t_i, \delta_i)$ and the j-th object $(\mathbf{x}_j, t_j, \delta_j)$ that can be compared. The weighting scheme is based on the Kaplan-Meier estimator, which considers the survival function defined in Equation (1). Based on the indicator r_{ij} defined earlier, the assigned weight is the survival function value of the j-th object.

To obtain the optimal parameter values, the Lagrangian function is constructed as follows:

$$L(\mathbf{w}, \xi; \alpha) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}C\sum_{i=1}^{n}\sum_{i<j}^{n} r_{ij}B_{ij}\xi_{ij}^2 - \sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{\alpha}_{ij}(\mathbf{w}^T\varphi(\mathbf{x}_j) - \mathbf{w}^T\varphi(\mathbf{x}_i) - 1 + \xi_{ij}) \tag{19}$$

where $\boldsymbol{\alpha}_{ij}$ are Lagrange multipliers. The Lagrangian function is differentiated with respect to $\mathbf{w}$, ξ, and α, yielding the following results:

$$\frac{\partial L_p}{\partial \mathbf{w}} = 0 \rightarrow \mathbf{w} - \sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij}(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i)) = 0 \rightarrow \mathbf{w} = \sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij}(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i)) \tag{20}$$

$$\frac{\partial L_p}{\partial \xi_{ij}} = 0 \rightarrow C\sum_{i=1}^{n}\sum_{i<j}^{n} r_{ij}B_{ij}\xi_{ij} - \sum_{i=1}^{n}\sum_{i<j}^{n}(-\alpha_{ij}) = 0 \rightarrow \alpha_{ij} = -Cr_{ij}B_{ij}\xi_{ij} \tag{21}$$

$$\begin{aligned}\frac{\partial L_p}{\partial \alpha} = 0 \quad &\rightarrow \quad -\sum_{i=1}^{n}\sum_{i<j}^{n}\left(\mathbf{w}^T(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i)) - 1 + \xi_{ij}\right) = 0 \\ &\rightarrow \sum_{i=1}^{n}\sum_{i<j}^{n}\left(1 + \xi_{ij} - \mathbf{w}^T(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i))\right) = 0\end{aligned} \tag{22}$$

The parameters $\mathbf{w}$ and ξ_{ij} are rewritten as follows:

$$\mathbf{w}^T = \left[\sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij}(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i))\right]^T \tag{23}$$

and

$$\xi_{ij} = -\frac{\alpha_{ij}}{Cr_{ij}B_{ij}} \tag{24}$$

Substituting Equations (23) and (24) into Equation (22) yields

$$\begin{aligned}&\sum_{i=1}^{n}\sum_{i<j}^{n}\left(1 + \xi_{ij} - \mathbf{w}^T(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i))\right) = 0 \\ &\Leftrightarrow \sum_{i=1}^{n}\sum_{i<j}^{n}1 - \sum_{i=1}^{n}\sum_{i<j}^{n}\frac{\alpha_{ij}}{Cr_{ij}B_{ij}} - \left\{\left[\sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij}\big(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i)\big)\right]^{\mathrm{T}}(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i))\right\} = 0\end{aligned} \tag{25}$$

If each term is multiplied by $Cr_{ij}B_{ij}$, we obtain:

$$\begin{aligned}&C\sum_{i=1}^{n}\sum_{i<j}^{n}r_{ij}B_{ij} - \sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij} - C\sum_{i=1}^{n}\sum_{i<j}^{n}r_{ij}B_{ij}\left\{\left[\sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij}\big(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i)\big)\right]^{\mathrm{T}}(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i))\right\} = 0 \\ &\Leftrightarrow C\sum_{i=1}^{n}\sum_{i<j}^{n}r_{ij}B_{ij} = \sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij} + C\sum_{i=1}^{n}\sum_{i<j}^{n}r_{ij}B_{ij}\left\{\left[\sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij}\big(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i)\big)\right]^{\mathrm{T}}(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i))\right\} = 0\end{aligned} \tag{26}$$

r_{ij} is an indicator of whether two individuals are comparable. To facilitate the computation of the prognostic index for each object, the expression is reformulated in matrix form. In Equation (25), the left-hand side can be expressed as in Equation (27).

$$C\sum_{i=1}^{n}\sum_{i<j}^{n}r_{ij}B_{ij} = C(r_{12}B_{12} + r_{13}B_{13} + \ldots + r_{(m-1)m}B_{(m-1)m}) = C\mathbf{B}_{(m\times m)}\mathbf{1}_{(m\times 1)} \tag{27}$$

where m is the number of comparable individuals. Next, the right-hand side involves summations with the identity

matrix so that the resulting matrix is square. For consistency, matrix dimensions must match. The first term on the right-hand side of Equation (25) is expressed as:

$$\sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \alpha_{12} \\ \alpha_{12} \\ \vdots \\ \alpha_{(n-2)n} \\ \alpha_{(n-1)n} \end{bmatrix} = \mathbf{I}_{(m\times m)}\alpha_{(m\times 1)} \tag{28}$$

The second term on the right-hand side of Equation (25) is given in Equation (29)

$$C\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} r_{ij}B_{ij} = Cr_{12}B_{12} + Cr_{13}B_{13} + \cdots + Cr_{(m-1)m}B_{(m-1)m}$$

$$\begin{aligned} &= C\begin{bmatrix} r_{12} & 0 & \cdots & 0 \\ 0 & r_{13} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{(m-1)m} \end{bmatrix} \begin{bmatrix} B_{12} & 0 & \cdots & 0 \\ 0 & B_{13} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_{(m-1)m} \end{bmatrix} \\ \Leftrightarrow\ & C\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} r_{ij}B_{ij} = C\begin{bmatrix} B_{12} & 0 & \cdots & 0 \\ 0 & B_{13} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_{(m-1)m} \end{bmatrix} = C\mathbf{B}_{(m\times m)} \end{aligned} \tag{29}$$

where $\mathbf{B}$ is a diagonal matrix with weight elements defined such that for each comparable pair of objects (t_i, δ_i, x_i) and (t_j, δ_j, x_j), B_{ij} is assigned a constant value for censored data ($\delta_j = 0$) and $B_{ij} = S(t_j)$ for event data ($\delta_j = 1$), where $S(t)$ is defined in Equation (1).

Several algebraic manipulations in matrix form are performed to obtain the prognostic index of each individual in the dataset. The matrix form of Equation (25) can be expressed as follows:

$$\begin{aligned} C\sum_{i=1}^{n}\sum_{i<j}^{n} r_{ij}B_{ij} = \sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij} + C\sum_{i=1}^{n}\sum_{i<j}^{n} r_{ij}B_{ij}\left\{\left[\sum_{i=1}^{n}\sum_{i<j}^{n}\alpha_{ij}\big(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i)\big)\right]^{\mathrm{T}}(\varphi(\mathbf{x}_j) - \varphi(\mathbf{x}_i))\right\} = 0 \\ \Leftrightarrow C\mathbf{B1} = \mathbf{I} + C\mathbf{B}(\alpha^T\mathbf{D}\mathbf{K}^T\mathbf{D}^T)^T \\ \Leftrightarrow C\mathbf{B1} = \mathbf{I}\alpha + C\mathbf{B}\mathbf{D}\mathbf{K}\mathbf{D}^T\boldsymbol{\alpha} \\ \Leftrightarrow C\mathbf{B}\mathbf{D}\mathbf{K}\mathbf{D}^T + \mathbf{I}\boldsymbol{\alpha} = C\mathbf{B1} \\ \Leftrightarrow (C\mathbf{B}\mathbf{D}\mathbf{K}\mathbf{D}^T + \mathbf{I})\boldsymbol{\alpha} = C\mathbf{B1} \\ \Leftrightarrow \boldsymbol{\alpha} = (C\mathbf{B}\mathbf{D}\mathbf{K}\mathbf{D}^T + \mathbf{I})^{-1}C\mathbf{B1} \end{aligned} \tag{30}$$

The matrix forms of $\mathbf{D}$, $\mathbf{K}$, and $\mathbf{B}$ used in Equation (30) are described as follows:

- The matrix $\mathbf{D}$ has elements $\{-1, 0, 1\}$, defined in Equation (11).

- The matrix $\mathbf{K}$ is an $n \times n$ kernel matrix where each element contains the kernel function value of the i-th and j-th individuals, given as $K(x_i, x_j) = \exp(-||x_i - x_j||^2/2\sigma^2)$.
- The weight matrix $\mathbf{B}$ is constructed based on the Kaplan-Meier estimator. For individuals with comparable survival times, the weights B_{ij} are constant for censored status and $B_{ij} = S(t_j)$ for event status, where $S(t)$ is defined in Equation (1).

In survival analysis, events occurring at longer times are typically less frequent. The model must assign greater importance to events since they carry crucial information for long-term survival estimation. The Kaplan-Meier-based weighting strategy for event status is designed to adjust the contribution of each event according to its survival probability up to time t.

For censored data, the exact event time is unknown; the only available information is that the subject survived up to time t, with no information on when the event will occur. Hence, this data provides partial information and its contribution to the model should not be ignored (weight not equal to 0). However, censored status should not be assigned a large weight due to the uncertainty of the event outcome.

The dimensions of the matrices used in Equation (30) are $\boldsymbol{\alpha}_{(m\times 1)}$, $\mathbf{B}_{(m\times m)}$, $\mathbf{D}_{(m\times n)}$, $\mathbf{K}_{(n\times n)}$, $\mathbf{I}_{(m\times m)}$, and $\mathbf{1}_{(m\times 1)}$.

The prognostic index is then defined as:

$$\begin{aligned}\hat{u}(x^*) &= \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\alpha}_{ij}(\phi(x_j) - \phi(x_i))^T\phi(x^*) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\alpha}_{ij}(\phi(x_j) - \phi(x_i))^T\phi(x^*) \\ &= \hat{\boldsymbol{\alpha}}^T\mathbf{D}\mathbf{K}(x^*) \end{aligned} \tag{31}$$

In general, the prognostic function of the Surv-WLSSVM model appears similar to the Surv-LSSVM model, with the distinction lying in the matrix $\boldsymbol{\alpha}$, which incorporates the weight matrix $\mathbf{B}$. The inclusion of weights in the Surv-WLSSVM model makes it more complex, specifically through the construction of the weight matrix, which in this study is derived from the Kaplan-Meier survival function.

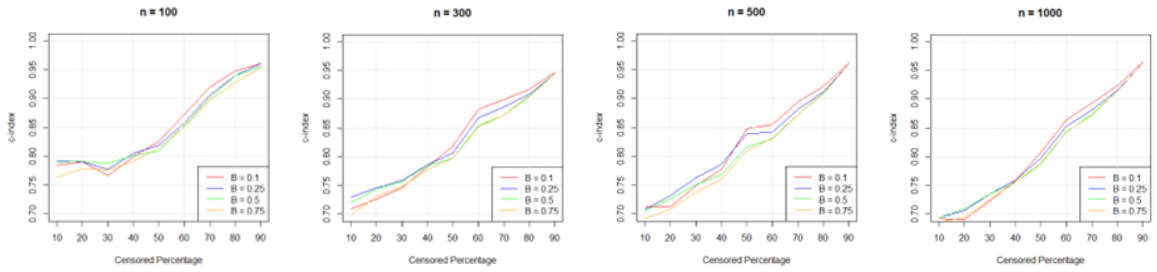
4.2. Model Implementation on Simulated Data

The simulated data used in this study consist of three datasets with 12, 17, and 35 variables, respectively. These datasets were generated from data following certain probability distributions. Survival times follow a Weibull distribution, censoring status is generated from random numbers 0 and 1, and predictor variables follow various distributions, including normal, Poisson, binomial, log-normal, and gamma. The performance of the Surv-WLSSVM model was evaluated on both small and large sample sizes, namely 100, 300, 500, and 1000.

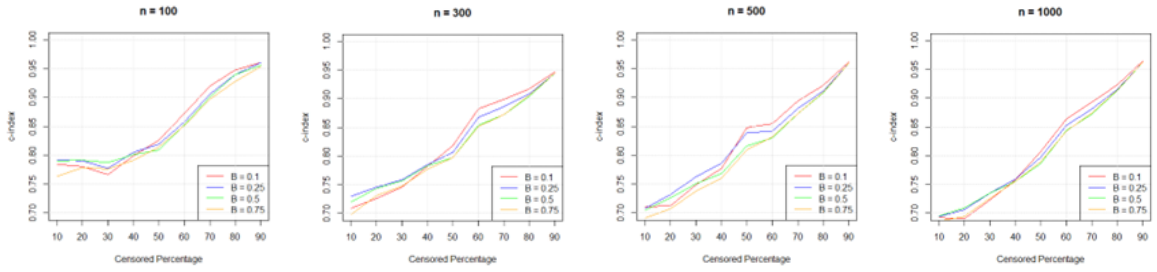
A comparison between the Surv-LSSVM and Surv-WLSSVM models was conducted to evaluate model performance with and without weighting, where weighting in this study was based on the Kaplan-Meier estimator.

The c-index value serves as the indicator of model goodness, applied to each dataset with censoring percentages ranging from 10%, 20%, up to 90% of the specified sample sizes. In survival analysis, a high level of censoring often poses a major challenge as it can reduce estimation efficiency and weaken statistical power. In general, a

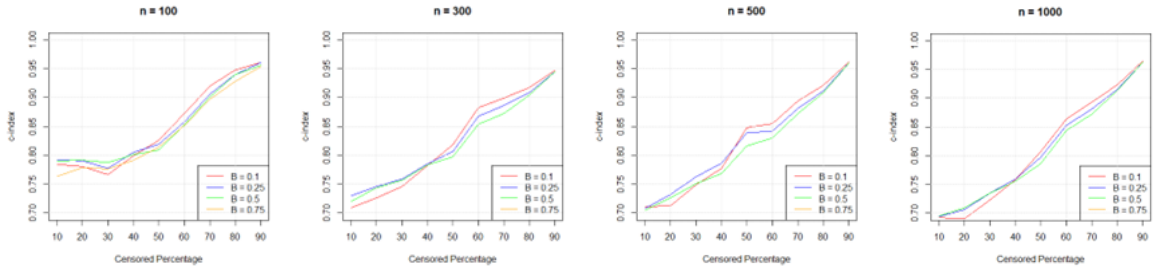
censoring proportion below 20% is considered low, while 20–40% is still regarded as common in medical studies. However, when the proportion of censored observations exceeds 50%, this condition is typically classified as high censoring, which may introduce bias and increase the uncertainty of the analytical results [68, 69]. Several simulation studies have even employed censoring scenarios of 60–70% to represent extreme situations that are frequently encountered in clinical data with limited follow-up periods. Weighting based on the Kaplan–Meier estimator accounts for the weights assigned to event data. Censored observations are assigned a constant weight; however, choosing an excessively large weight reduces the contribution of event observations, which contain complete survival information. Conversely, assigning a very small or zero weight causes the model to ignore information from censored data. The selection of the censored-data weight was performed by evaluating several candidate values that represent high, moderate, and low weighting levels. The results of this simulation study are presented in Figure 3.



(a) 12 variables



(b) 17 variables



(c) 35 variables

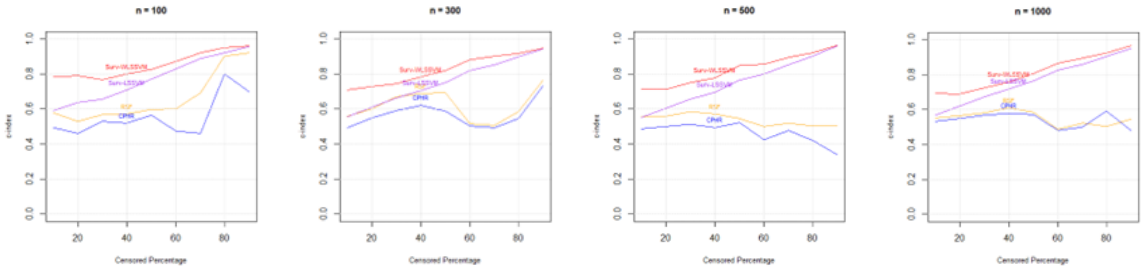
Figure 3. c-index with censored data weighting

Figure 3 presents a comparison of the performance of the Surv-WLSSVM model under various weighting schemes for censored data. The model employs the Kaplan-Meier survival function $S(t)$ as the weight for observed event data and combines it with four constant weights for censored observations, namely of 0.1, 0.25, 0.5, and 0.75. The analysis results indicate that using smaller weights for censored data, particularly $B = 0.1$, tends to yield slightly higher c-index values compared to other weighting schemes across most scenarios. The performance differences among the four weighting schemes were not statistically significant, suggesting that variations in the constant

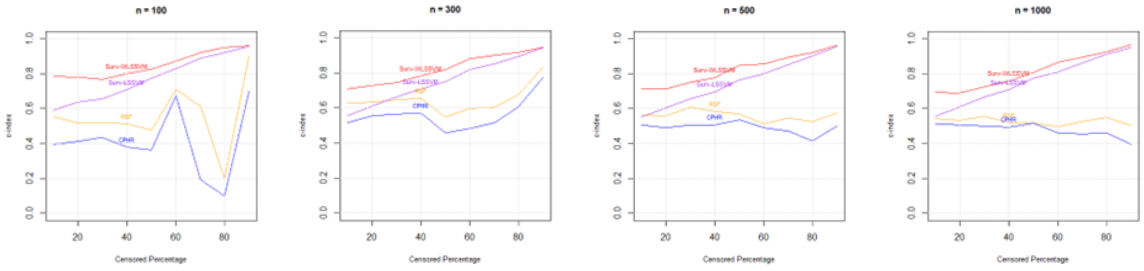
weight assigned to censored data do not substantially affect model accuracy. The Kaplan–Meier–based weighting strategy with varying constant weights produces comparable outcomes, indicating that the choice of constant weight for censored data can be adjusted flexibly without compromising the predictive quality of the model.

Parameter tuning in the model was performed using the Particle Swarm Optimization (PSO) algorithm, which searches for the optimal parameter values within the range [0.001 ; 1]. The application of the model across different sample sizes shows a consistent increase in c-index values from datasets with low to high censoring percentages.

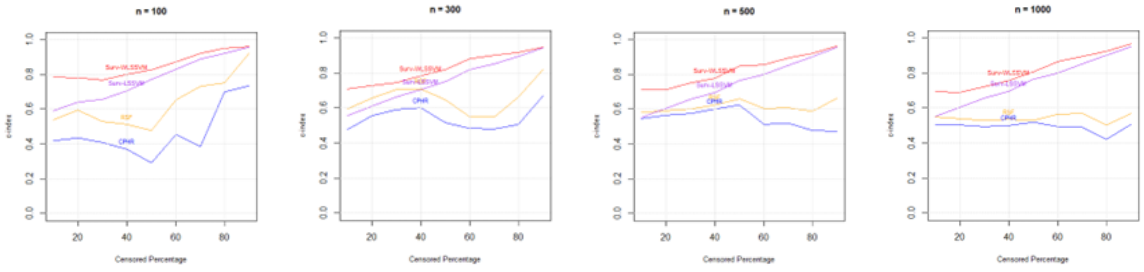
The performance of the Surv-WLSSVM model compared with Surv-LSSVM, RSF, and CPHR is presented in Figure 4.



(a) 12 variables



(b) 17 variables



(c) 35 variables

Figure 4. Comparison of c-index for Surv-WLSSVM, Surv-LSSVM, RSF, and CPHR models

Figure 4 presents a comparison of the concordance index (c-index) values for four survival analysis methods: Surv-WLSSVM, Surv-LSSVM, RSF, and CPHR, evaluated across different numbers of variables (12, 17, and 35) and sample sizes ($n = 100$, 300, 500, and 1000). The weighting scheme applied in the Surv-WLSSVM model uses the survival function value for observations that experienced an event, while a constant weight of $B = 0.1$ is assigned to censored data.

In general, the Surv-WLSSVM model demonstrates more stable and consistent performance compared to the other methods, particularly at higher censoring levels. Its c-index values tend to increase with larger sample sizes and are relatively less affected by the proportion of censored data. In contrast, the Surv-LSSVM, RSF, and CPHR models exhibit greater fluctuations, especially in small samples ($n = 100$) and under high censoring conditions, indicating higher sensitivity to information loss due to censoring. These results suggest that incorporating censored data weighting based on the survival function improves the estimation accuracy of the relationship between predictor variables and event time.

4.3. Model Implementation on Real Data

In addition to the implementation conducted on simulated data, model evaluation was also carried out on real datasets to measure the performance of the Surv-LSSVM and Surv-WLSSVM methods. The testing was performed on three datasets, namely Breastfeeding, PBC, and Bone Marrow. Each method was evaluated using the values of parameters C and σ optimized through a PSO-based parameter tuning process, with the main performance indicator being the c-index value, which represents the model's accuracy in ranking event risks. Figure 5 provides a description of the percentage of censored and event occurrences in the real datasets used in this study.

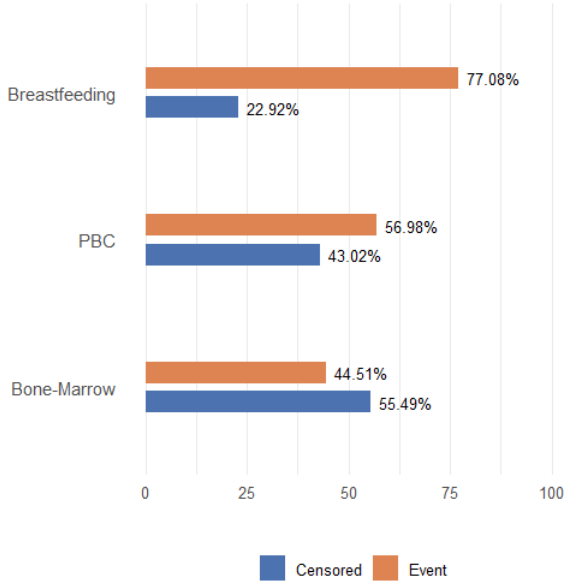


Figure 5. Percentage of censored data in real datasets

Figure 5 presents the distribution of the percentage between censored data and event occurrences across the three datasets: Breastfeeding, PBC, and Bone Marrow. It can be seen that the Breastfeeding dataset has the highest event proportion, at 77.08%, while only 22.92% of the data are censored. This indicates that most subjects in this dataset experienced the observed event during the follow-up period. Similarly, the PBC dataset shows that 56.98% of the data experienced an event, which is higher compared to the censored data of 43.02%.

The Bone Marrow dataset exhibits a relatively balanced distribution, with 55.49% censored data and 44.51% events. This pattern reflects a comparable challenge in survival data processing, particularly in ensuring that the model is not biased toward the dominance of censored data. The variation in proportions between censored and event data across the three datasets highlights the necessity of selecting survival analysis methods capable of

accommodating data imbalance and being flexible with event-time distributions.

Table 3 presents the performance of the Surv-WLSSVM, Surv-LSSVM, RSF, and CPHR models across three real datasets. Higher c-index values indicate superior model performance.

Table 3. Comparison of c-index Across Three Datasets

Model	Breastfeeding	PBC	Bone-Marrow
CPHR	0.03241650	0.22850679	0.28282828
RSF	0.22327682	0.13651884	0.50925932
Surv-LSSVM	0.73542757	0.77187466	0.82844666
Surv-WLSSVM	0.75200235	0.78948431	0.83803277

Both the CPHR and RSF models exhibit relatively low performance across the three datasets, particularly for the Breastfeeding and PBC datasets, indicating that linear modeling or tree-based partitioning alone is insufficient to capture the nonlinear structure and heterogeneity present in the survival data. The Surv-LSSVM model provides a substantial improvement in performance for all datasets, achieving c-index values of 0.735 for Breastfeeding, 0.772 for PBC, and 0.828 for Bone-Marrow. The kernel-based approach used in Surv-LSSVM allows the model to effectively capture nonlinear relationships between covariates and event risk—an aspect that cannot be optimally addressed by the semi-parametric CPHR model. Moreover, the stable performance of Surv-LSSVM across datasets with differing characteristics demonstrates its superior ability to handle variations in censoring rates, sample sizes, and the complexity of covariate interactions.

The incorporation of a weighting scheme in Surv-WLSSVM yields the best performance across all datasets, with c-index values of 0.752 (Breastfeeding), 0.789 (PBC), and 0.838 (Bone-Marrow). Weighting based on the Kaplan–Meier survival function enables the model to reduce the influence of censored observations and adapt to imbalances in information. These results show that integrating kernel methods with censoring-aware weighting consistently enhances survival prediction accuracy compared to both the semi-parametric CPHR model and the modern ensemble-based RSF model.

Censoring-based weighting may be considered to further improve Surv-WLSSVM performance. Assigning constant weights to all censored observations is less optimal, and a sensitivity analysis is required to assess the impact of weight selection on model performance. The use of constant weights does not adequately reflect the differing levels of information contributed by early-censored versus late-censored observations.

This study employed the PSO algorithm to determine the optimal parameters. The tendencies of the regularization parameter, kernel parameter, and c-index for each dataset are presented sequentially in Figure 6

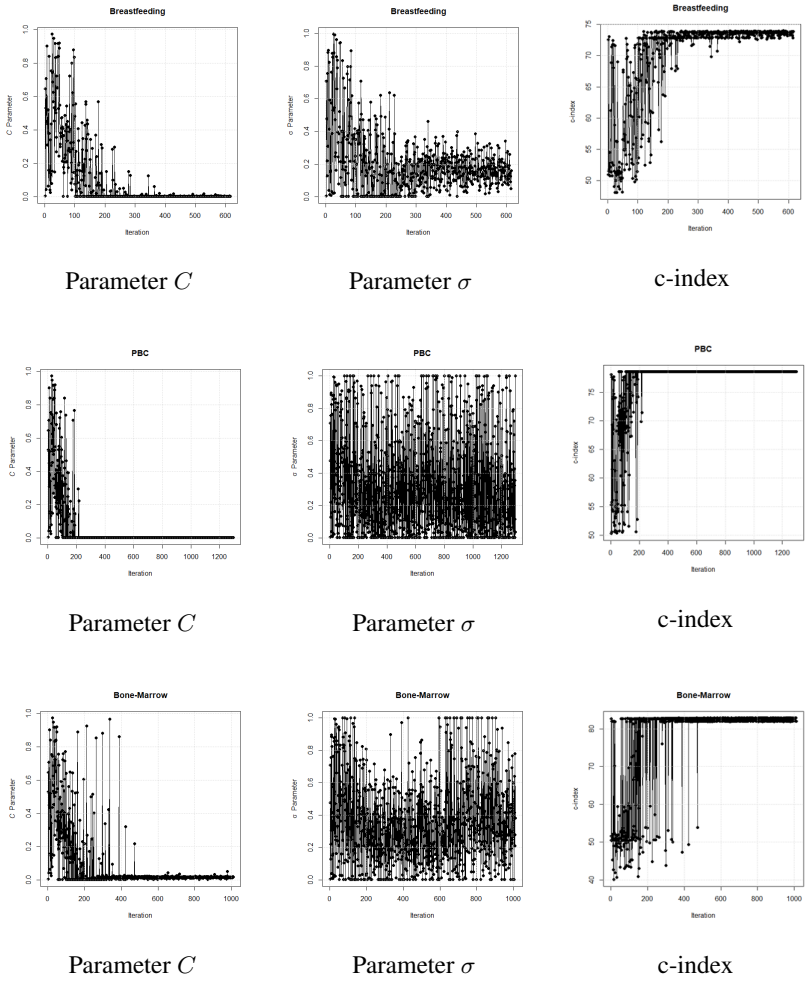


Figure 6. The Parameter Values of the Surv-WLSSVM Model for the Real Datasets

Based on Figures 6, the visualization results show relatively consistent patterns across all datasets. The parameter C tends to converge more quickly within a certain range, indicating the model's stability against variations in C once the optimal regularization threshold is reached. In other words, once the model obtains a sufficient C value to control complexity and regularization, further changes in C do not provide significant improvements in performance. In contrast, the parameter σ exhibits more fluctuating and sensitive patterns, suggesting that kernel selection strongly influences the model's ability to capture the nonlinear structure of the data.

In the Breastfeeding dataset, the c-index reaches its maximum value at a relatively sharp parameter combination, particularly with respect to changes in σ, while C shows stability after passing a certain point. A similar pattern is observed in the PBC dataset, with broader fluctuations along the σ axis, indicating that more complex data structures require more precise kernel adjustments. In the Bone Marrow dataset, although the area with high c-index values appears broader, the fluctuations in σ remain dominant compared to C, again highlighting the rapid convergence tendency of the regularization parameter.

These findings confirm that the parameter C is more stable and easier to regulate through the optimization process, while σ requires broader exploration and higher precision due to its high sensitivity to the data distribution. Therefore, the use of PSO-based optimization algorithms is highly effective in addressing complex and non-convex

parameter dynamics as found in the Surv-WLSSVM model. This approach enables adaptation to the characteristics of each dataset and contributes to improving the accuracy and generalization of the SVM-based survival analysis model.

Weighting within the framework of Survival Least Squares Support Vector Machine is effective in enhancing the accuracy of survival analysis models, particularly in more complex data situations. The fundamental difference between Surv-LSSVM and Surv-WLSSVM lies in the objective function used and how each method handles the characteristics of survival data. Surv-LSSVM treats all data equally without considering weighting. In contrast, Surv-WLSSVM employs a weighting approach that enables the model to adjust the contribution of each data pair based on its significance or risk level.

Regarding sensitivity to censored data, Surv-LSSVM tends to be less responsive because it does not directly account for censoring variation during the learning process. On the other hand, Surv-WLSSVM is more sensitive to data since the weighting mechanism can be adjusted, allowing pairs with stronger information or higher risk to receive greater weights, while weaker or less informative pairs can have their influence minimized.

In terms of flexibility, Surv-LSSVM exhibits limitations because it applies uniform weights to all observation pairs, making it less capable of capturing variations in data structure. In contrast, Surv-WLSSVM is more adaptive to complex data through its weighting mechanism, which helps reduce the influence of weak pairs and enhances the stability of risk ranking. This enables Surv-WLSSVM to deliver generally more consistent performance, particularly in datasets with high censoring levels, data imbalance, or heterogeneous risk patterns.

Nevertheless, the performance of Surv-WLSSVM based on the Kaplan–Meier estimator still holds substantial potential for improvement. The use of constant weights for censored observations does not distinguish between early-censored and late-censored cases, even though these conditions convey different levels of information in survival analysis. By incorporating a more appropriate weighting scheme for censored data, the performance of Surv-WLSSVM can be further enhanced.

Table 4. Differences Between Surv-LSSVM and Surv-WLSSVM

Aspect	Surv-LSSVM	Surv-WLSSVM
Objective function	Squared loss without weights	Squared loss with weights
Sensitivity to data	Less attentive	More sensitive due to adjustable weights
Flexibility	Less flexible in handling data pairs	Highly flexible, adaptive to data structure
Risk ranking accuracy	Can be affected by weak pairs	More stable
Computational efficiency	Faster without weighting operations	Slower due to weight-matrix construction

The weighting mechanism in terms of computational efficiency indicates that Surv-LSSVM operates faster because it does not require additional weighting procedures, resulting in a simpler computational process. In contrast, Surv-WLSSVM, which employs the Kaplan–Meier estimator, tends to be slower due to the involvement of more complex formulas and the need to construct a weight matrix that scales with the sample size. Interestingly, several previous studies have shown that the WLSSVM classification method in general contexts actually has lower computational

complexity compared to LSSVM. Nevertheless, within the framework of survival analysis, the implementation of Surv-WLSSVM becomes more computationally expensive due to the requirement for a more structured weighting scheme.

The Surv-WLSSVM approach offers several advantages that provide added value compared with classical survival analysis techniques and previous SVM-based methods. First, this model integrates a weighting strategy into the Surv-LSSVM framework, enabling it to account for differences in the contributions of event and censored units. Second, Surv-WLSSVM retains the strengths of the LSSVM formulation, namely its quadratic optimization structure, which can be solved more efficiently.

On the other hand, the Surv-WLSSVM method also has several limitations. First, the model's performance strongly depends on the selection of hyperparameters, such as the regularization parameter and kernel parameters, thereby requiring an optimization process that can be computationally demanding, particularly when population-based algorithms such as PSO are employed. Second, although the weighting strategy improves accuracy, the approach remains sensitive to extreme weights, especially in datasets with a very small proportion of events, a condition that may lead to an imbalance in contributions across observations. Third, the model assumes a non-informative censoring mechanism; violations of this assumption may degrade model performance.

5. Discussion

The Surv-WLSSVM and Surv-LSSVM models have been thoroughly evaluated through simulations and applications on real-world data. The simulation design was constructed with predictor complexities of 12, 17, and 35 variables, as well as varying sample sizes of 100, 300, 500, and 1000. In addition to variations in dimension and sample size, differences in censoring proportions were also simulated to assess the robustness of the models across different levels of censoring. The results demonstrated that Surv-WLSSVM consistently outperformed across all simulation scenarios.

This superiority is driven by the integration of survival function weights derived from the Kaplan-Meier estimator. This value is only available for subjects who experienced the event, assigning proportional weights based on actual risk. Conversely, censored subjects do not have survival values and are assigned constant weights. When the weight for censored data, Surv-WLSSVM exhibited a significant improvement in c-index values compared to Surv-LSSVM, particularly in scenarios with a large proportion of events. On the other hand, in scenarios with high percentages of censoring, although Surv-WLSSVM remained numerically superior, the differences were not statistically significant. In such cases, the large number of censored data and uniform weighting of this type of data resulted in no difference in contribution among censored subjects, so the model only distinguished predictor variables through kernel mapping.

This condition emphasizes the importance of adaptive weighting strategies for censored data. If partial information from censored subjects can be further utilized in determining weights, the potential for enhancing model performance could be substantially greater.

The application of the Surv-WLSSVM and Surv-LSSVM models to real-world data indicates that the differences in c-index values are not statistically significant. This condition may be attributed to the distributional characteristics of the real dataset, both in terms of survival times and predictor variables. In contrast, the RSF and CPHR models produce substantially lower c-index values. The kernel-based approach employed in Surv-WLSSVM and Surv-LSSVM provides greater flexibility for capturing nonlinear patterns that cannot be accommodated by the CPHR model. Meanwhile, the RSF model is relatively less responsive to censored observations and requires careful parameter tuning to improve its performance.

The Surv-WLSSVM model demonstrates strong performance across both small scale datasets and big data environments. For small datasets such as clinical health screening data (e.g., hypertension or stroke assessments) or household electronic device lifespan data the method provides stable estimates due to the efficiency of the LSSVM formulation under limited sample sizes. Moreover, Surv-WLSSVM aligns well with modern analytical demands in the era of big data, as it is capable of handling large numbers of variables and observations, including genomic expression data, daily vehicle sensor records, and digital application usage histories. The consistently strong performance of Surv-WLSSVM across all simulated data scenarios highlights its relevance for a wide range of real-life datasets that are familiar in daily life and require accurate predictions of durability or time to event outcomes.

The implementation of a survival-function–based weighting strategy contributes meaningfully to enhancing the model's sensitivity to subject-specific characteristics, while also opening opportunities for further methodological development through adaptive weighting schemes for censored data. This study recommends using the Kaplan–Meier Imputation with Bayesian approach for weighting censored observations, or alternatively employing weighting based on Inverse Probability of Censoring Weighting (IPCW).

6. Conclusion

The findings of this study demonstrate that the Surv-WLSSVM model consistently delivers superior performance compared with Surv-LSSVM, RSF, and CPHR across a wide range of testing conditions. In the simulated datasets, this superiority was evident in all scenarios, including variations in the number of predictors (data dimensionality), sample size, and censoring proportions. The performance gap between Surv-WLSSVM and Surv-LSSVM became statistically significant under low-censoring conditions, where both constant weighting and Kaplan–Meier–based weighting contributed substantially to the improvement of Surv-WLSSVM. Under conditions with a high proportion of events, both SVM-based models achieved c-index values exceeding 0.90. In contrast, the RSF and CPHR models exhibited less stable c-index values across simulation scenarios; RSF requires optimal parameter-search procedures to improve performance, while the semi-parametric CPHR model performs suboptimally when confronted with nonlinear structures and high-dimensional predictor spaces.

Applications to real-world datasets further confirmed the superior performance of Surv-WLSSVM. The three datasets featured different combinations of censored and event observations, yet Surv-WLSSVM consistently maintained high predictive accuracy. The model's performance is strongly influenced by the selection of the regularization and kernel parameters, both of which are essential components of SVM-based methods. The parameter tuning procedure using the PSO algorithm yielded more stable and well-regulated regularization parameters, whereas the kernel parameter required broader search-space exploration to attain optimal values.

The methodological integration of survival analysis, weighted least squares support vector machines, the Kaplan–Meier estimator, and PSO-based parameter tuning constitutes a robust analytical framework capable of delivering superior performance in modern survival analysis settings. The Surv-WLSSVM model represents a survival modeling approach with substantial potential for more reliable and adaptive future development.

References

[1] S. Deng, Y. Zhou, J. Lin, Z. Xiang, L. Liu, N. Xie, H. Yu, and Y. Sun. Survival Analysis of Gynecological Cancers in Southeast China, 2011–2020: A Population-Based Study. *Cancer Epidemiology*, 92(July):102641, 2024.

[2] L. H. Al Fryan and M. B. Alazzam. Survival Analysis of Oncological Patients Using Machine Learning Method. *Healthcare*, 11(1), 2023.

[3] A. S. Azar, S. B. Rikan, A. Naemi, J. B. Mohasefi, H. Pirnejad, M. B. Mohasefi, and U. K. Wiil. Application of Machine Learning Techniques for Predicting Survival in Ovarian Cancer. *BMC Medical Informatics and Decision Making*, 22(1):1–24, 2022.

[4] J. L. Tonetto, J. M. Pique, A. Fochezatto, and C. Rapetti. Survival Analysis of Small Business during COVID-19 Pandemic, a Brazilian Case Study. *Economies*, 12(7), 2024.

[5] D. Xu. Quantitative modeling of M&A Success Probability: Integrating Macroeconomic Volatility and Temporal Factors through Survival Analysis. *Journal of Economic Technology*, 3(January):202–222, 2025.

[6] A. Saha, L. Hock-Eam, and S. G. Yeok. Drivers of MSME Loan Defaults Using Survival Analysis: Implications for Lenders and Policy Planners. *International Journal of Emerging Markets*, 20(7):2816–2836, 2025.

[7] W. Tizi and A. Berrado. Machine Learning for Survival Analysis in Cancer Research: A Comparative Study. *Scientific African*, 21(December 2022):e01880, 2023.

[8] D. Papathanasiou, K. Demertzis, and N. Tziritas. Machine Failure Prediction Using Survival Analysis. *Future Internet*, 15(5), 2023.

[9] S. Germer, C. Rudolph, L. Labohm, A. Katalinic, N. Rath, K. Rausch, B. Holleczek, and H. Handels. Survival Analysis for Lung Cancer Patients: A Comparison of Cox Regression and Machine Learning Models. *International Journal of Medical Informatics*, 191(March), 2024.

[10] Y. Jia, C. Li, C. Feng, S. Sun, Y. Cai, P. Yao, X. Wei, Z. Feng, Y. Liu, W. Lv, H. Wu, F. Wu, L. Zhang, S. Zhang, and X. Ma. Prognostic Prediction for Inflammatory Breast Cancer Patients Using Random Survival Forest Modeling. *Translational Oncology*, 52(September 2024):102246, 2025.

[11] C. Y. Bae, B. S. Kim, S. H. Jee, J. H. Lee, and N. D. Nguyen. A Study on Survival Analysis Methods Using Neural Network to Prevent Cancers. *Cancers*, 15(19), 2023.

[12] C. T. Ho, E. C. Tan, P. C. Lee, C. J. Chu, Y. H. Huang, T. I. Huo, M. C. Hou, J. C. Wu, and C. W. Su. Prognostic Nutritional Index as a Prognostic Factor for Very Early-Stage Hepatocellular Carcinoma. *Clinical and Translational Gastroenterology*, 15(4):1–8, 2024.

[13] Y. Deng, L. Liu, H. Jiang, Y. Peng, Y. Wei, Z. Zhou, Y. Zhong, Y. Zhao, X. Yang, J. Yu, Z. Lu, A. Kho, H. Ning, N. B. Allen, J. T. Wilkins, K. Liu, D. M. Lloyd-Jones, and L. Zhao. Comparison of State-of-the-Art Neural Network Survival Models with the Pooled Cohort Equations for Cardiovascular Disease Risk Prediction,.

[14] C. G. Meixide and M. Matabuena and L. Abraham and M. R. Kosorok. Neural interval-censored survival regression with feature selection. *Statistical Analysis and Data Mining*, 17(4):1–12, 2024.

[15] J. Kiessling and A. Brunnberg and G. Holte and N. Eldrup and K. Sörelius. Artificial Intelligence Outperforms Kaplan–Meier Analyses Estimating Survival after Elective Treatment of Abdominal Aortic Aneurysms. *European Journal of Vascular and Endovascular Surgery*, 65(4):600–607, 2023.

[16] J. Shen, J. Wu, M. Xu, D. Gan, B. An, and F. Liu. A Hybrid Method to Predict Postoperative Survival of Lung Cancer Using Improved SMOTE and Adaptive SVM. *Computational and Mathematical Methods in Medicine*, 2021, 2021.

[17] H. Sanz, F. Reverter, and C. Valim. Enhancing SVM for Survival Data Using Local Invariances and Weighting. 21(193):1–20, 2020.

[18] V. Van Belle and K. Pelckmans and S. Van Huffel and J. A. K. Suykens. Support vector methods for survival analysis: A comparison between ranking and regression approaches. *Artificial Intelligence in Medicine*, 53(2):107–118, 2011.

[19] C. Nguyen and P. Huynh. Weighted Least Square - Support Vector Machine. In *Proceedings - 2021 RIVF International Conference on Computing and Communication Technologies*, pages 1–6, 2021.

[20] I. El Badisy, Z. BenBrahim, M. Khalis, S. Elansari, Y. ElHitmi, F. Abbass, N. Mellas, and K. El Rhazi. Risk Factors Affecting Patients Survival with Colorectal Cancer in Morocco: Survival Analysis Using an Interpretable Machine Learning Approach. *Scientific Reports*, 14(1):1–13, 2024.

[21] S. W. Lee. Kaplan-Meier and Cox Proportional Hazards Regression in Survival Analysis: Statistical Standard and Guideline of Life Cycle Committee. *Life Cycle*, 3:1–9, 2023.

[22] M. I. Akazue, G. A. Nwokolo, O. A. Ejaita, C. O. Ogeh, and E. Ufiofio. Machine Learning Survival Analysis Model for Diabetes Mellitus. *International Journal of Innovative Science, Research & Technology*, 8(4):754–760, 2023.

[23] R. Sujatha and J. M. Chatterjee and N. Z. Jhanjhi and T. A. Tabbakh and Z. A. Almusaylim. Heart failure patient survival analysis with multi kernel support vector machine. *Intelligent Automation & Soft Computing*, 32(1):115–129, 2022.

[24] S. Leurcharusmee, P. Maneejuk, W. Yamaka, N. Thaiprasert, and N. Tuntichiranon. Survival Analysis of Thai Micro and Small Enterprises During the Covid-19 Pandemic. *Journal of Business Economics and Management*, 23(5):1211–1233, 2022.

[25] J. Suykens V. Van Belle, K. Pelckmans and S. Van Huffel. Additive survival least squares support vector machines. *Statistics in Medicine*, 29(2):296–317, 2010.

[26] C. Khotimah, S. W. Purnami, D. D. Prastyo, V. Chosuvivatwong, and H. Sriplung. Additive survival least square support vector machines: A simulation study and its application to cervical cancer prediction. *AIP Conference Proceedings*, 1908(1):050013, 2017.

[27] H. A. Khoiri, D. D. Prastyo, and S. W. Purnami. A simulation study and application of feature selection on survival least square support vector machines. *AIP Conference Proceedings*, 2138(1):050016, 2019.

[28] N. G. Mikhaeel, M. W. Heymans, J. J. Eertink, H. C. W. de Vet, R. Boellaard, U. Duhrsen, L. Ceriani, C. Schmitz, S. E. Wiegers, A. Huttmann, P. J. Lugtenburg, E. Zucca, G. J. C. Zwezerijnen, O. S. Hoekstra, J. M. Zijlstra, and S. F. Barrington. Proposed New Dynamic Prognostic Index for Diffuse Large B-Cell Lymphoma: International Metabolic Prognostic Index. *Journal of Clinical Oncology*, 40(21):2352–2360, 2022.

[29] D. G. Kleinbaum. *Survival Analysis Self Learning*. Springer, London, third edition, 2012.

[30] K. Jóźwiak, V. H. Nguyen, L. Sollfrank, S. C. Linn, and M. Hauptmann. Cox Proportional Hazards Regression in Small Studies of Predictive Biomarkers. *Scientific Reports*, 14(1), December 2024.

[31] P. Saha, Y. Marouf, H. Pozzebon, A. Guergachi, K. Keshavjee, M. Noaeen, and Z. Shakeri. Predicting Time to Diabetes Diagnosis Using Random Survival Forests. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*. IEEE, 2024.

[32] A. Garcia-Lopez, M. Jiménez-Gómez, A. Gomez-Montero, J. C. Gonzalez-Sierra, S. Cabas, and F. Giron-Luque. Survival Analysis Using Machine Learning in Transplantation: A Practical Introduction. *BMC Medical Informatics and Decision Making*, 25(1), December 2025.

[33] O. Ben-Assuli, R. Ramon-Gonen, T. Heart, A. Jacobi, and R. Klempfner. Utilizing Shared Frailty with the Cox Proportional Hazards Regression: Post Discharge Survival Analysis of CHF Patients. *Journal of Biomedical Informatics*, 140, April 2023.

[34] W. Tizi and A. Berrado. Machine Learning for Survival Analysis in Cancer Research: A Comparative Study. *Scientific African*, 21, September 2023.

[35] A. AlShehhi, T. M. Almansoori, A. R. Alsuwaidi, and H. Alblooshi. Utilizing Machine Learning for Survival Analysis to Identify Risk Factors for COVID-19 Intensive Care Unit Admission: A Retrospective Cohort Study from the United Arab Emirates. *PLoS One*, 19(1 January), January 2024.

[36] Q. Xianxin, J. Gao, J. Yang, J. Hu, W. Hu, L. Kong, and J. J. Lu. A Comparison Study of Machine Learning (Random Survival Forest) and Classic Statistic (Cox Proportional Hazards) for Predicting Progression in High-Grade Glioma after Proton and Carbon Ion Radiotherapy. *Frontiers in Oncology*, 10:1–10, 2020.

[37] M. Farhadian, S. Dehdar Karsidani, A. Mozayanimonfared, and H. Mahjub. Risk Factors Associated with Major Adverse Cardiac and Cerebrovascular Events Following Percutaneous Coronary Intervention: A 10-year follow-up Comparing Random Survival Forest and Cox proportional-Hazards Model. *BMC Cardiovascular Disorders*, 21:1–8, 2021.

[38] M. Sharafi et al. Factors Associated with Survival of Tuberculosis Patients in Southeast Iran: Comparison of Stepwise Cox Regression, Survival Tree, and Random Survival Forest. *BMC Public Health*, 25(1), December 2025.

[39] T. Liao, T. Su, Y. Lu, L. Huang, W. Wei, and L. H. Feng. Random Survival Forest Algorithm for Risk Stratification and Survival Prediction in Gastric Neuroendocrine Neoplasms. *Scientific Reports*, 14(1), December 2024.

[40] A. Saeed et al. Random Survival Forest Model for Early Prediction of Alzheimer's Disease Conversion in Early and Late Mild Cognitive Impairment Stages. *PLoS One*, 19(12 December), December 2024.

[41] M. Sharafi, M. A. Mohsenpour, S. Afrashteh, M. H. Eftekhari, A. Dehghan, A. Farhadi, A. Jafarnezhad, A. Zakeri, and M. Azizmohammad Looha. Factors Affecting the Survival of Prediabetic Patients: Comparison of Cox Proportional Hazards Model and Random Survival Forest Method. *BMC Medical Informatics and Decision Making*, 24(1), December 2024.

[42] V. H. Truong and H. A. Pham. Support Vector Machine for Regression of Ultimate Strength of Trusses: A Comparative Study. *Engineering Journal*, 25(7):157–166, 2021.

[43] Z. Jun. The Development and Application of Support Vector Machine. *Journal of Physics: Conference Series*, 1748(5), 2021.

[44] F. Khan, Y. Gulzar, S. Ayoub, M. Majid, M. S. Mir, and A. B. Soomro. Least square-support vector machine based brain tumor classification system with multi model texture features. *Frontiers in Applied Mathematics and Statistics*, 9, 2023.

[45] R. Wang, Y. Sun, J. Ni, and H. Zheng. Identification of Product Definition Patterns in Mass Customization by Multi-Information Fusion Weighted Support Vector Machine. *Engineering Applications of Artificial Intelligence*, 137(PB):109253, 2024.

[46] K. P. Widiatmika. Fault Diagnosis of PV Array Based on Time Series and Support Vector Machine. *Etika Jurnalisme Pada Koran Kuning Sebuah Studi Mengenai Koran Lampu Hijau*, 16(2):39–55, 2015.

[47] X. Yang, Q. Song, and Y. Wang. A Weighted Support Vector Machine for Data Classification. *International Journal of Pattern Recognition and Artificial Intelligence*, 21(5):961–976, 2007.

[48] S. Heddam, M. Ptak, M. Sojka, S. Kim, A. Malik, O. Kisi, and M. Zounemat-Kermani. Least Square Support Vector Machine-Based Variational Mode Decomposition: A New Hybrid Model for Daily River Water Temperature Modeling. *Environmental Science and Pollution Research*, 29(47):71555–71582, 2022.

[49] Z. Gholami, M. H. Ahmadi Azqhandi, M. Hosseini Sabzevari, and F. Khazali. Evaluation of Least Square Support Vector Machine, Generalized Regression Neural Network and Response Surface Methodology in Modeling the Removal of Levofloxacin and Ciprofloxacin from Aqueous Solutions Using Ionic Liquid @Graphene oxide@ ionic liquid NC. *Alexandria Engineering Journal*, 73:593–606, 2023.

[50] H. J. Jang, S. J. Shin, and A. Artemiou. Principal Weighted Least Square Support Vector Machine: An Online Dimension-Reduction Tool for Binary Classification. *Computational Statistics & Data Analysis*, 187:107818, 2023.

[51] J. A. K. Suykens and J. Vandewalle. Least squares support vector machine classifiers. *Kluwer Academic Publishers*, pages 1–10, 1999.

[52] A. H. Poursaeed and F. Namdari. Real-time Voltage Stability Monitoring Using Weighted Least Square Support Vector Machine Considering Overcurrent Protection. *International Journal of Electrical Power & Energy Systems*, 136(September 2021):107690, 2022.

[53] A. Y. A. Amer. Global-Local Least-Squares Support Vector Machine (GLocal-LS-SVM). *PLoS ONE*, 18(4 April):1–18, 2023.

[54] J. A. K. Suykens, J. De Brabanter, L. Lukas, and J. Vandewalle. Weighted Least Squares Support Vector Machines: Robustness and Sparce Approximation. *Neurocomputing*, 48(1-4):85–105, 2002.

[55] K. L. Du, B. Jiang, J. Lu, J. Hua, and M. N. S. Swamy. Exploring Kernel Machines and Support Vector Machines: Principles, Techniques, and Future Directions. *Mathematics*, 12(24):1–58, 2024.

[56] M. Daviran, A. Maghsoudi, and R. Ghezelbash. Optimized AI-MPM: Application of PSO for Tuning the Hyperparameters of SVM and RF Algorithms. *Computers & Geosciences*, 195(November 2024):105785, 2025.

[57] P. Royston and D. G. Altman. External Validation of a Cox Prognostic Model: Principles and Methods. *BMC Medical Research Methodology*, 13(1), 2013.

[58] A. F. Ferreira, T. Fernandes, M. do C. Carvalho, and H. S. Loureiro. The Prognostic Role of Prognostic Nutritional Index and Controlling Nutritional Status in Predicting Survival in Older Adults with Oncological Disease: A Systematic Review. *Onco*, 4(2):101–115, 2024.

[59] V. Van Belle, K. Pelckmans, J. A. K. Suykens, and S. Van Huffel. Learning Transformation Models for Ranking and Survival Analysis. *Journal of Machine Learning Research*, 12:819–862, 2011.

[60] M. A. Abido. Optimal power flow using particle swarm optimization. *International Journal of Electrical Power and Energy Systems*, 24(7):563–571, October 2002.

[61] Z. Gaing. Particle Swarm Optimization to Solving the Economic Dispatch Considering the Generator Constraints. 18:1187–1195, February 2014.

[62] K. Zhang, Y. Xu, W. Gao, K. Yu, and Y. Liu. Optimizing MRT Data Processing via Comparative Analysis of SA and PSO Algorithms: a Simulation and Numerical Study. *Optics Express*, 31(13):20905, 2023.

[63] G. Papazoglou and P. Biskas. Review and Comparison of Genetic Algorithm and Particle Swarm Optimization in the Optimal Power Flow Problem. *Energies*, 16(3), 2023.

[64] K. Rosen, V. Prasad, and E. Y. Chen. Censored Patients in Kaplan–Meier Plots of Cancer Drugs: An Empirical Analysis of Data Sharing. *European Journal of Cancer*, 141:152–161, 2020.

[65] T. Dwi Tama and E. Astutik. Exclusive Breastfeeding Survival And Factors Related to Early Breastfeeding Cessation in Indonesia. January 2019.

[66] E. R. Dickson, P. M. Grambsch, T. R. Fleming, L. D. Fisher, and A. Langworthy. Prognosis in Primary Biliary Cirrhosis: Model for Decision Making. *Hepatology*, 10(1):1–7, 1989.

[67] M. Kwon, B. S. Kim, S. Yoon, S. O. Oh, and D. Lee. Hematopoietic Stem Cells and Their Niche in Bone Marrow. *International Journal of Molecular Sciences*, 25(13), 2024.

[68] J. P. Klein and M. L. Moeschberger. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer, 2 edition, 2003.

[69] D. W. Hosmer, S. Lemeshow, and S. May. *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*. Wiley, 2 edition, 2008.