Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: www.tandfonline.com/journals/tstf20

An AUC-based multi-kernel weighted support vector machine ensemble algorithm for breast cancer diagnosis

Mushuang Cheng, Lintong Liu, Haixiang Lin & Guoqiang Wang

To cite this article: Mushuang Cheng, Lintong Liu, Haixiang Lin & Guoqiang Wang (22 Dec 2025): An AUC-based multi-kernel weighted support vector machine ensemble algorithm for breast cancer diagnosis, Statistical Theory and Related Fields, DOI: 10.1080/24754269.2025.2603548

To link to this article: https://doi.org/10.1080/24754269.2025.2603548

Published online: 22 Dec 2025.

STATISTICAL THEORY AND RELATED FIELDS
https://doi.org/10.1080/24754269.2025.2603548

An AUC-based multi-kernel weighted support vector machine ensemble algorithm for breast cancer diagnosis

Mushuang Cheng[a], Lintong Liu[a], Haixiang Lin[b,c] and Guoqiang Wang [a]

[a]School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai, People's Republic of China; [b]Delft Institute of Applied Mathematics, Delft University of Technology, Delft, The Netherlands; [c]Institute of Environmental Sciences (CML), Leiden University, Leiden, The Netherlands

ABSTRACT

Machine learning algorithms have demonstrated outstanding performance for disease diagnosis. Kernel function selection plays a crucial role in effectively transforming the nonlinear nature of input data. To enhance breast cancer diagnosis, we propose a novel ensemble algorithm, namely, AUC-Ada-L_1MKL-WSVM, which integrates Weighted Support Vector Machines (WSVM), AdaBoost, and Multi-Kernel Learning (MKL). This ensemble algorithm introduces two main innovations. First, it simultaneously updates the weights of training samples and the combined kernel function during classification. Second, it incorporates an AUC-based approach to adjust training sample weights, effectively controlling the growth rate of misclassified sample weights in AdaBoost. Experimental results are provided to demonstrate the effectiveness of our method, which achieves an AUC score of 97.21% and an accuracy of 97.64% on the WDBC dataset, and an AUC of 97.53% and an accuracy of 97.46% on the WBC dataset. Comparative analysis further confirms that our ensemble algorithm outperforms four benchmark models in classification accuracy.

ARTICLE HISTORY
Received 24 May 2025
Revised 20 November 2025
Accepted 9 December 2025

KEYWORDS
Weighted support vector machine; breast cancer diagnosis; ensemble algorithm; multi-kernel learning; AdaBoost

1. Introduction

According to the data from the World Health Organization in 2020, breast cancer has become the most prevalent cancer worldwide, thereby posing a significant threat to women's health. Because its pathogenesis is not well understood, many patients remain unaware of their condition until the disease has progressed to an advanced stage. Missing the optimal treatment window puts lives at risk. Therefore, early detection of breast cancer is crucial for effective medical treatment and disease control. However, even for experienced doctors, distinguishing between malignant and benign breast tumors using traditional diagnostic techniques remains challenging. In response, machine learning has emerged as a powerful tool in the biomedical field, demonstrating significant potential in early breast cancer detection. Consequently, researchers are actively developing machine learning-based methods for early breast cancer diagnosis to assist doctors in tumor classification (Aymaz, 2025; Ghani et al., 2019; Jha et al., 2017).

CONTACT Guoqiang Wang guoq_wang@hotmail.com School of Mathematics, Physics and Statistics, Shanghai University of Engineering Science, Shanghai 201620, People's Republic of China

Among various approaches, machine learning algorithms based on image analysis, such as mammograms, histological images, and ultrasound imaging have been enhanced by neural networks, leading to significant breakthroughs in breast cancer diagnosis (Aljuaid et al., 2022; Barnett et al., 2021; J. L. Li et al., 2017; Ragab et al., 2021). Additionally, some diagnostic methods rely on features extracted from these images. Two well-known examples are the Wisconsin Diagnostic Breast Cancer (WDBC) and Wisconsin Breast Cancer (WBC) datasets. In this paper, we focus on machine learning algorithms that leverage these extracted features to enhance diagnostic accuracy.

SVM and its extended models have been widely used in disease diagnosis due to their strong generalization ability (Ali et al., 2024; Sahu & Mohanty, 2021; Sharma et al., 2021; J. J. Wang et al., 2023). For SVM, the kernel function is a critical factor influencing classification performance (Le & Clarke, 2022). Many studies have focussed on traditional single-kernel learning, but MKL has been shown to significantly improve the interpretability of SVM models (Bach et al., 2004). The classification performance of SVM largely depends on the selection of kernel functions and sample weight allocation. To address these challenge, Rakotomamonjy and Grandvalet (2008) proposed the MKL method, which enhances SVM adaptability by automatically selecting optimal kernel functions. Meanwhile, some studies have attempted to improve SVM classification performance through sample weighting strategies. For example, X. C. Li et al. (2008) proposed an AdaBoost-based SVM approach that incorporates SVM into the resampling stage, thereby improving classification accuracy.

In recent years, researchers have further explored the combination of SVM kernel function optimization and weighting strategies. Ramirez-Morales et al. (2023) employed a Genetic Algorithm for automatic kernel selection to improve the generalization ability of SVM. F. Wang et al. (2019) enhanced AdaBoost-SVM by incorporating feature learning and proposed a training framework based on kernel parameter tuning to improve classification performance. Luo et al. (2020) introduced a weighted SVM ensemble model based on an AdaBoost-SVM structure, incorporating sample weighting to handle imbalanced datasets.

However, existing methods still have two major limitations. First, most studies focus solely on kernel function optimization or sample weighting, without effectively integrating the advantages of both approaches. For instance, Xie et al. (2022) proposed a regression model that combines these two strategies, but its application in classification tasks remains unexplored. Second, AdaBoost increases the weights of hard-to-classify samples during training. However, these weights tend to grow exponentially, leading to severe weight imbalance among training samples. This imbalance adversely affects model stability and generalization capability.

To address these challenges, this paper introduces AUC-Ada-L_1MKL-WSVM, a multi-kernel weighted SVM ensemble algorithm designed for breast cancer diagnosis. This method integrates sample reweighting strategies into the MKL framework, thereby improving model robustness in classification tasks. Furthermore, we develop an AUC-based AdaBoost weight update strategy to mitigate sample weight imbalance during training. Experimental results confirm that AUC-Ada-L_1MKL-WSVM surpasses existing methods in breast cancer classification, demonstrating its practical performance. Our main contributions are as follows.

(1) We propose a novel ensemble algorithm for early breast cancer diagnosis, which integrates MKL with an enhanced sample weighting strategy. This unified approach simultaneously exploits kernel optimization and adaptive weighting to improve classification performance.

(2) We introduce a refined sample weighting mechanism based on AUC, which mitigates the issue of weight imbalance in AdaBoost, leading to more stable and generalizable models.
(3) Extensive empirical evaluations on benchmark breast cancer datasets demonstrate that our algorithm achieves higher AUC and classification accuracy than conventional methods.

The paper is structured as follows. Section 2 reviews SVM, their extended models, ensemble algorithms for breast-cancer diagnosis, and core concepts of machine-learning classification. Section 3 details the methodology of the study and the evaluation metrics used to gauge performance. Section 4 introduces the proposed ensemble algorithm, AUC-Ada-L_1MKL-WSVM, which subsumes L_1MKL-SVM, L_1MKL-WSVM, and Ada-L_1MKL-WSVM as special cases. Section 5 summarizes the experimental setup and results. Section 6 concludes the study and outlines future research directions.

2. Literature review

This section reviews SVM and its extensions in breast cancer diagnosis, highlighting their advantages over traditional laboratory tests. Additionally, recent ensemble algorithms for breast cancer diagnosis are also discussed.

2.1. SVM and its extended models for breast cancer diagnosis

Research on single-kernel SVM has gained traction, yielding significant advancements in breast cancer diagnosis. For example, Asri et al. (2016) compared SVM against K nearest neighbour, Naive Bayes, and Decision tree classifiers using the WBC dataset, with SVM outperforming the rest in terms of accuracy. Akay (2009) proposed a model that combines feature selection with SVM. Experimental results indicated that SVM with five selected features achieved an accuracy of 99.51% on the WBC dataset. Chen et al. (2011) proposed a model combining rough set theory with SVM, where a rough set reduction algorithm was used for feature selection. The effectiveness of this model has been confirmed on the WBC dataset. Zheng et al. (2014) presented a model that combines k-means and SVM. The model obtains an accuracy of 97.38% on the WDBC dataset. In addition, some researchers have investigated extended variants of SVM for breast cancer diagnosis. Polat and Güneş (2007) employed least squares SVM for breast cancer diagnosis. The model achieved an accuracy of 98.53%. Liu et al. (2019) utilized a cost-sensitive SVM model combined with feature selection for classification. However, a single kernel function cannot effectively transform the nonlinear nature of input data into a more expressive feature space. Some researchers have conducted studies on multi-kernel SVM in response to this challenge. For example, Sannasi Chakravarthy et al. (2022) presented a hybrid model of mixture kernel SVM based on ebola optimization algorithm and applied it to the WDBC dataset.

Though SVM-based models have shown efficacy in breast cancer diagnosis, their performance is highly dependent on the fine-tuning of critical parameters, such as the kernel type and associated hyperparameters. In this study, we focus on optimizing the selection of these key parameters.

2.2. Ensemble algorithms for breast cancer diagnosis

Three well-known basic ensemble algorithms are Bagging, Boosting and Stacking. Their purpose is to reduce the bias and variance of classification models. Several studies have explored the application of these ensemble techniques in breast cancer diagnosis (Kapila & Saleti, 2023). For example, Abdar et al. (2020) proposed a two-layer nested ensemble model, in which the second-layer meta-classifier incorporates two or three different algorithms. Experimental results on the WDBC dataset showed that this model outperformed individual base models. Nanglia et al. (2022) developed a heterogeneous ensemble model that combines k-nearest neighbours, SVMs, and decision trees for breast cancer diagnosis.

In ensemble learning, the use of ROC curves instead of accuracy has gained increasing attention as a more comprehensive performance metric (Gao et al., 2006; Levesque et al., 2012). Gao and Sun (2007) demonstrated that an AUC-based learning strategy can improve the effectiveness of conventional model parameter tuning. Nonetheless, most existing approaches focus on adjusting model parameters based on ROC curves, rather than modifying the importance of training samples accordingly. As an improvement, H. F. Wang et al. (2018) proposed an ensemble model that integrates twelve pre-tuned SVMs, using AUC-based criteria to guide the combination process. Experimental results show that this ensemble algorithm reduces variance and enhances accuracy in breast cancer diagnosis. However, it is noteworthy that these studies do not explore how ROC curves can be leveraged to adjust the importance of training instances.

3. Related methods

In this section, we first introduce two models related to SVM, namely Weighted Support Vector Machine (WSVM) and Boost-SVM. These two models provide us with the main ideas of our ensemble algorithm. Secondly, we introduce some evaluation indicators for classification models used in this article.

3.1. WSVM

In SVM, the regularization coefficient is identical for all samples. However, the importance of individual samples may vary significantly. WSVM, in contrast to standard SVM, assigns different weights to training samples to address class imbalance and varying sample importance. Each sample should be assigned a weight based on its importance. In practice, it is desirable that important samples are correctly classified. For less significant samples, it is acceptable to tolerate their misclassification (Lin & Wang, 2002). Yang et al. (2005) extended SVM by introducing sample-specific weight factors into the regularization term, resulting in the WSVM.

Let $S = \{(x_i, y_i), x_i \in X, y_i \in \{-1, 1\}, i = 1, \ldots, N\}$ be a training dataset, where X is a n-dimensional feature space. Then WSVM is given by

$$\begin{aligned} \min \quad & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} \lambda_i \xi_i \\ \text{s.t.} \quad & \begin{cases} y_i(\langle w, \phi(x_i)\rangle + b) \geq 1 - \xi_i, & i = 1, 2, \ldots, N, \\ \xi_i \geq 0, & i = 1, 2, \ldots, N, \end{cases} \end{aligned} \tag{1}$$

Table 1. Kernel functions and default parameters settings.

Kernel type	Functions	Default parameters
Linear	$\kappa(x_i, x_j) = \langle x_i, x_j \rangle$	
Polynomial	$\kappa(x_i, x_j) = (\langle x_i, x_j \rangle + 1)^d$	$d = 1$
Gaussian	$\kappa(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{\sigma^2})$	$\sigma = 1$
Sigmoid	$\kappa(x_i, x_j) = \tanh(\langle x_i, x_j \rangle + b)$	$b = 1$

where C is the regularization coefficient (larger C allocates higher penalties for misclassification), ξ_i is the slack variable, b is the intercept, λ_i is the weight of the training sample, and $\phi(x_i)$ is the map function. Furthermore, the commonly used kernel functions and their default parameter settings are presented in Table 1.

In our ensemble algorithm, we adopt the idea of WSVM, which considers the importance of the samples. Detailed explanations can be found in Yang et al. (2005).

3.2. Boost-SVM

Compared with SVM, WSVM assigns different weights to training samples to handle class imbalance and variations in sample importance. However, WSVM still suffers from limited performance improvement when facing complex data distributions or noisy instances, as the weight adjustment alone may not sufficiently adapt the model. To address this limitation, ensemble methods have been introduced. Zhang and Ren (Zhang & Ren, 2008) proposed an ensemble algorithm, Boost-SVM, which integrates SVM into the AdaBoost framework. This approach enhances the classification performance of SVM by iteratively adjusting sample weights during training. Detailed explanations can be found in Kashef (2021).

The flow of Boost-SVM is given in Algorithm 1. In our ensemble algorithm, we also adopt the resampling idea of Boost-SVM. In addition, we have discovered the limitation of updating rules in Boost-SVM, which provides us with room for improvement.

3.3. Evaluation indicators

It is well known that many evaluation metrics are derived from the confusion matrix, which consists of two classes, labelled as positive $(+1)$ and negative (-1). In this study, we utilize three evaluation metrics based on the confusion matrix: accuracy (ACC), sensitivity (SE), and specificity (SP). Additionally, we employ the AUC as an evaluation metric. A higher AUC value indicates a stronger classification ability of the model. Therefore, AUC is used to further assess the performance of our proposed algorithm. For more details on the ROC curve and AUC, please refer to relevant literature (H. F. Wang et al., 2018).

4. The proposed ensemble algorithm

In this section, we introduce a new ensemble algorithm, AUC-Ada-L_1MKL-WSVM. The algorithm is developed in the following steps. First, we incorporate the idea of MKL by using the L_1 norm to combine multiple base kernels into a composite kernel for SVM. This allows adaptive selection of the optimal base kernels and their parameters, resulting in L_1MKL-SVM. Next, we extend L_1 MKL-SVM by introducing sample weighting factors, forming L_1MKL-WSVM. The kernel function weights are solved using the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA). We then integrate L_1MKL-WSVM with FISTA

Algorithm 1 Boost-SVM Algorithm

Input: Training dataset $S = \{(x_i, y_i)\}_{i=1}^{N}$, where x_i represents the feature vector and $y_i \in \{-1, 1\}$ denotes the class label.

Output: Final ensemble classifier: $F(x)$.

1: **Initialize:** Set initial sample weights $D_i = \frac{1}{N}, \forall i \in \{1, \ldots, N\}$.
2: **for** $t = 1$ to T **do**
3: Sample training instances according to the distribution D_i^t and train an SVM classifier.
4: Obtain the weak classifier: $g_t(x) : X \to \{-1, 1\}$.
5: Compute the classification error: $\text{error}_t = \sum_{i=1}^{N} D_i^t I(g_t(x_i) \neq y_i)$.
6: Compute the weight of the classifier:

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \text{error}_t}{\text{error}_t} \right). \tag{2}$$

7: Update the sample weights:

$$D_i^{t+1} = \frac{D_i^t \exp(-y_i \alpha_t g_t(x_i))}{Z_t}, \quad \forall\, i \in \{1, \ldots, N\}. \tag{3}$$

8: Normalize D_i^{t+1} using the normalization constant Z_t.
9: **end for**
10: **Return:** Final ensemble classifier:

$$F(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t g_t(x) \right). \tag{4}$$

into AdaBoost to update sample weights dynamically, resulting in Ada-L_1MKL-WSVM. Finally, to address weight distortion caused by AdaBoost, we propose AUC-Ada-L_1MKL-WSVM, which enhances weight adaptation and classification performance.

While WSVM, MKL, and AdaBoost are well-established methods, our approach aims to integrate them in a unified framework rather than simply combining them. The main novelty is the use of an AUC-guided joint update for both kernel weights and sample weights. This design allows the model to adapt at two levels simultaneously, which is different from existing MKL or AdaBoost frameworks. By doing so, the method reduces the negative influence of noisy or minority samples and improves robustness in imbalanced medical data, such as breast cancer diagnosis.

4.1. L_1MKL-SVM

MKL improves the predictive capacity of the model to a moderate extent by incorporating the unique characteristics of different kernel functions through a weighted combination of these basic kernel functions.

A commonly acknowledged fact is that the use of the L_1 norm results in sparse weights for the combined kernel function, and it can enhance the computational efficiency of the model. Thus, we introduce the L_1 norm into the objective function of SVM. This yields the so-called L_1MKL-SVM, i.e.,

$$\begin{aligned} \min \quad & \frac{1}{2}\left(\sum_{q=1}^{Q}\|w_q\|)^2\right) + C\sum_{i=1}^{N}\xi_i + \gamma\|D\|_1 \\ \text{s.t.} \quad & \begin{cases} \sum_{q=1}^{Q} y_i[\langle w_q, \sqrt{d_q}\phi_q(x_i)\rangle + b] \geq 1-\xi_i, \quad i=1,2,\ldots,N, \\ \xi_i \geq 0, \quad i=1,2,\ldots,N, \\ d_q \geq 0, \quad q=1,2,\ldots,Q, \end{cases} \end{aligned} \tag{5}$$

where γ is the regularization coefficient, and $D = \{d_q \mid q = 1, 2, \ldots, Q\}$ are the weights of the combined kernel function. Since the product of d_q and w_q is nonconvex, we apply the variable transformation $w'_q = \sqrt{d_q}w_q$ to reformulate (5) as

$$\begin{aligned} \min \quad & \frac{1}{2}\sum_{q=1}^{Q}\frac{\|w'_q\|^2}{d_q} + C\sum_{i=1}^{N}\xi_i + \gamma\|D\|_1 \\ \text{s.t.} \quad & \begin{cases} \sum_{q=1}^{Q} y_i[\langle w'_q, \phi_q(x_i)\rangle + b] \geq 1-\xi_i, \quad i=1,2,\ldots,N, \\ \xi_i \geq 0, \quad i=1,2,\ldots,N, \\ d_q \geq 0, \quad q=1,2,\ldots,Q. \end{cases} \end{aligned} \tag{6}$$

It should be noted that (6) can be solved via convex optimization algorithm (Xie et al., 2022).

4.2. L_1MKL-WSVM

In order to also consider the importance of each training sample, we add the weights λ of the training samples to L_1MKL-SVM. This yields the so-called L_1MKL-WSVM, i.e.,

$$\begin{aligned} \min \quad & \frac{1}{2}\sum_{q=1}^{Q}\frac{\|w'_q\|^2}{d_q} + C\sum_{i=1}^{N}\lambda_i\xi_i + \gamma\|D\|_1 \\ \text{s.t.} \quad & \begin{cases} \sum_{q=1}^{Q} y_i[\langle w'_q, \phi_q(x_i)\rangle + b] \geq 1-\xi_i, \quad i=1,2,\ldots,N, \\ \xi_i \geq 0, \quad i=1,2,\ldots,N, \\ d_q \geq 0, \quad q=1,2,\ldots,Q. \end{cases} \end{aligned} \tag{7}$$

When $\lambda_i = 1$, L_1MKL-WSVM degenerates to L_1MKL-SVM.

Let

$$M(D) = \frac{1}{2}\sum_{q=1}^{Q}\frac{\|w'_q\|^2}{\widetilde{d}_q} + C\sum_{i=1}^{N}\lambda_i\widetilde{\xi}_i \tag{8}$$

and $(\widetilde{b}, \widetilde{\xi}, \widetilde{w'})$ be an optimal solution of

$$\begin{aligned} \min \quad & \frac{1}{2}\sum_{q=1}^{Q}\frac{\|w'_q\|^2}{d_q} + C\sum_{i=1}^{N}\lambda_i\xi_i \\ \text{s.t.} \quad & \begin{cases} \sum_{q=1}^{Q} y_i[\langle w'_q, \phi_q(x_i)\rangle + b] \geq 1 - \xi_i, \quad i = 1, 2, \ldots, N, \\ \xi_i \geq 0, \quad i = 1, 2, \ldots, N, \\ d_q \geq 0, \quad q = 1, 2, \ldots, Q. \end{cases} \end{aligned} \tag{9}$$

It follows from (1) that

$$M(D) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j \sum_{q=1}^{Q} d_q\kappa_q(x_i, x_j) + \sum_{i=1}^{N}\alpha_i, \tag{10}$$

where α_i is the Lagrange multiplier, and $\kappa(x_i, x_j) = \langle\phi(x_i), \phi(x_j)\rangle$ is the kernel function.

Then, we can reformulate (7) as below.

$$\min Z(D) = M(D) + \gamma\|D\|_1, \quad D \geq 0. \tag{11}$$

One can easily verify that (11) is a compound objective optimization problem. Furthermore, due to the non-differentiability of $\|D\|_1$, we can utilize the notion of the proximal gradient (Nesterov, 2013) to solve it. In this work, FISTA is employed to update D (Beck & Teboulle, 2009).

Furthermore, we can obtain the gradient of $M(D)$ from

$$\nabla M(d_q^{(t-1)}) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j\kappa_q(x_i, x_j). \tag{12}$$

For more details please refer to Xie et al. (2022).

4.3. Ada-L_1MKL-WSVM

To concurrently update the weights of the training samples, we embed FISTA into Algorithm 1, and an ensemble algorithm is established, namely Ada-L_1MKL-WSVM.

The steps of Ada-L_1MKL-WSVM are as follows.

Step 1: Construct a basic classifier by training a group of training samples with MKL-SVM.

Step 2: Calculate the classification error rate of the basic classifier from (2).

Step 3: Update the new weights D_i^{t+1} of each training subset from (3).

Step 4: Obtain a new training subset by resampling the classification samples from D_i^{t+1}.

Step 5: Calculate the weights of the combined kernel function on the new training subset.

Step 6: Obtain a strong classifier by combining the basic classifiers acquired during each iteration.

Table 2. Sensitivity analysis of the quantile parameter τ on the WDBC dataset.

τ	AUC	ACC	SP	SE
0.1	0.9680 ± 0.0026	0.9730 ± 0.0022	0.9880 ± 0.0012	0.9250 ± 0.0052
0.2	0.9690 ± 0.0025	0.9740 ± 0.0021	0.9890 ± 0.0014	0.9400 ± 0.0052
0.3	$\mathbf{0.9714 \pm 0.0028}$	$\mathbf{0.9759 \pm 0.0023}$	$\mathbf{0.9915 \pm 0.0012}$	$\mathbf{0.9512 \pm 0.0054}$
0.4	0.9694 ± 0.0028	0.9743 ± 0.0024	0.9912 ± 0.0015	0.9402 ± 0.0053
0.5	0.9685 ± 0.0025	0.9735 ± 0.0021	0.9860 ± 0.0015	0.9450 ± 0.0052

4.4. AUC-Ada-L_1MKL-WSVM

In Ada-L_1MKL-WSVM, the weights of the training samples are updated based on the classification error rate of each weak classifier, as referred to in (3). Given that the classification error rate is constrained to be less than 0.5, Ada-L_1MKL-WSVM inherently focuses more attention on training samples that are incorrectly classified. Notably, when a training sample is consistently misclassified across iterations, the weight assigned to that sample increases exponentially, potentially leading to serious distortion in the sample weights.

To mitigate the aforementioned issue, we introduce an enhancement to Ada-L_1MKL-WSVM by combining AUC to update sample weights. Unlike the error rate, which only accounts for the correctness of classification, AUC provides a probabilistic evaluation of a sample being assigned to a particular class and is particularly advantageous when dealing with imbalanced class distributions. In this case, AUC offers a more discriminating metric than error rate.

The core innovation involves the substitution of

$$\alpha_t = \tau * \mathrm{AUC}_t \tag{13}$$

with (2), where AUC_t is the AUC value of the t-th weak classifier.

Since the quantile parameter τ plays a central role in shaping the asymmetric margin of the proposed model, we further performed a sensitivity analysis to ensure reproducibility and fair comparison. Using the same 10-fold cross-validation splits, we evaluated $\tau \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ on the WDBC dataset and reported ACC, AUC, SE, and SP as mean $\pm$ SD over the ten folds.

As shown in Table 2, $\tau = 0.3$ achieves the highest ACC (0.9759 ± 0.0023) and AUC (0.9714 ± 0.0028), together with the best sensitivity–specificity balance. These results indicate that the asymmetric margin induced by $\tau = 0.3$ aligns well with the underlying class distribution, and therefore we adopt $\tau = 0.3$ as the default setting in all subsequent experiments.

In this article, we choose $\tau = 0.3$. It should be noted that if the current base classifier's AUC is greater than 0.5 or hasn't reached the maximum number of iterations, the next iteration will be performed; otherwise, the next iteration will not proceed.

The variation range of α_t in (13) is smaller than that in (2). This can effectively control the speed of the weight increase for erroneous samples in the next iteration and reduce the weight difference between samples in two adjacent iterations.

This refinement leads to our proposed ensemble algorithm, named AUC-Ada-L_1MKL-WSVM. The flow chart is shown in Algorithm 2. And in Algorithm 2, η represents the learning rate for updating kernel weights, $k^{(1)}$ specifies the number of initial kernels chosen at the beginning of the optimization, and tol denotes the tolerance threshold that serves as the stopping criterion for convergence.

Algorithm 2 AUC-Ada-L_1MKL-WSVM

Input: Training samples S;

$L^{(0)} = l\,(l \geq 1)$, $\eta = 2$, $k^{(1)} = 1$, C, γ, tol, τ.

Output: The strong classifier: $F(x) = \text{sign}(\sum_{t=1}^{T} \alpha_t g_t(x))$.

Initialize: $D^{(0)} = (d_1^{(0)}, \ldots, d_Q^{(0)}) = (1/Q, \ldots, 1/Q)$.

Initialize $D_i^t = (w_1, w_2, \ldots, w_N) = (\frac{1}{N}, \frac{1}{N}, \ldots, \frac{1}{N})$.

$H^{(1)} = D^{(0)}$.

for $t = 1$ to $\ldots$ **do**

    Train MKL-SVM based on D_i^t, and obtain a classifier $g_t(x)$.

    Calculate the AUC value of $g_t(x)$.

    if AUC $> \frac{1}{2}$ **then**

        $\alpha_t = \tau \cdot \text{AUC}_t$.

    else

        break.

    end if

    Optimize D through FISTA. See Xie et al. (2022).

end for

Given the importance of Algorithm 2, we provide a more detailed textual description of its workflow to complement the pseudocode. The algorithm starts by initializing kernel and sample weights, and then iteratively trains MKL-SVM classifiers. In each iteration, the AUC value of the weak classifier is calculated to determine the update coefficient α_t. If the classifier achieves an AUC greater than 0.5, the sample weights are updated accordingly; otherwise, the iteration is terminated. Meanwhile, the kernel weights are optimized using FISTA until convergence under the tolerance criterion. This process is repeated until the maximum number of iterations is reached, and all weak classifiers are finally combined to form the strong classifier.

In this framework, AUC is adopted as the guiding metric because it highlights the ranking of high-risk (malignant) samples, which is of practical importance in early diagnosis. Although AUC does not directly reflect the confusion matrix, it provides a probabilistic evaluation particularly suitable for imbalanced medical data. At the same time, we acknowledge that other measures, such as the Matthews Correlation Coefficient (MCC) (Chicco & Jurman, 2020), can also provide valuable insights into classification performance. As part of future work, we plan to investigate MCC-based weight updating as a complementary strategy to further enhance model stability and accuracy.

5. Experiment and results

5.1. Dataset description

In this paper, we choose the WDBC [1] and WBC [2] datasets which are freely accessible in the UCI repository. The overview of each dataset is given in Table 3. Sixteen samples containing missing values of attributes in the WBC dataset are deleted.

[1] https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic.

[2] https://archive.ics.uci.edu/dataset/15/breast+cancer+wisconsin+original.

Table 3. Dataset description.

Dataset	Characteristic number	Sample size	Missing value
WDBC	30	569	0
WBC	9	699	16

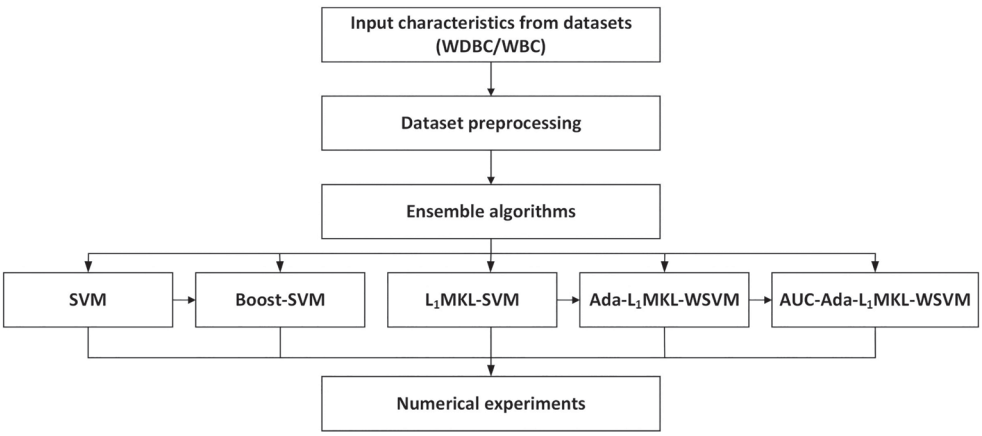

Figure 1. The workflow of our experiment AUC-Ada-L1MKL-WSVM.

For the WDBC dataset, 357 samples are malignant and 212 are benign, while for the WBC dataset, 241 samples are malignant and 458 are benign after removing 16 instances with missing values. Although the imbalance ratio is moderate in both datasets, it still poses challenges for detecting malignant cases. Therefore, the use of weighted SVM can help emphasize these critical but relatively fewer malignant instances, which is important for improving early cancer diagnosis.

5.2. Data preprocessing

In order to eliminate the impact of dimensionality, we use the max-min normalization method to normalize the WDBC and WBC datasets, i.e.,

$$\tilde{a}_{i,j} = \frac{a_{i,j} - a_i^{\min}}{a_i^{\max} - a_i^{\min}}, \tag{14}$$

where $a_{i,j}$ is the j-th value corresponding to the i-th feature, and $a_i^{\max}$ and $a_i^{\min}$ are the maximun and minimum values of the i-th feature, respectively.

5.3. Workflow

The workflow of our experiment is shown in Figure 1. The experiment in this paper is performed on a personal computer with Intel(R) Core(TM) i7-12700K @ 3.60GHz and 32GB memory. The operating system is Windows 10 and the program is implemented in Python 3.10.9.

Table 4. Parameter settings and definitions.

Parameter	Value	Definition
C	[1, 10, 100]	SVM regularization parameter
σ	[100, 50, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.005, 0.001]	Gaussian kernel parameter
d	[1, 2]	Polynomial kernel degree
β	0.009	FISTA stopping tolerance
l	100	FISTA step size
b	1	Sigmoid kernel bias
γ	[0.01, 0.1, 1]	Regularization parameter

5.4. Parameter settings

For SVM related parameters, based on the characteristics of the dataset, the range of values for C in this experiment is set to [1,10,100]. The parameters σ, d, and b are the main adjustable parameters of the Gaussian kernel function, Polynomial kernel function, and Sigmoid kernel function of SVM, respectively. For the Sigmoid kernel function, we choose the Sigmoid kernel function with its default parameters. For Gaussian kernel functions, we choose ten Gaussian kernel functions with values of 100, 50, 5, 1, 0.5, 0.1, 0.05, 0.001, 0.005 and 0.001, respectively. For polynomial kernel functions, we choose two polynomial kernel functions with being 1 and 2, respectively.

Regarding the parameters of FISTA, based on the characteristics of the dataset, for the stopping criteria, the tolerance value tol is set to 0.009, which also notifies the convergence. The maximum number of iterations is set to 7, and the step size of the internal gradient that controls the convergence speed is set to 100.

For the regularization parameter γ, we set its values to [0.01, 0.1, 1]. For each combination of C and γ, we use the grid search method to select the best combination.

To sum up, we selected ten Gaussian kernel functions with varying parameters, two polynomial kernel functions with varying parameters, and one Sigmoid kernel function. The parameter settings for our experiment are shown in Table 4.

The parameter settings used in this study were selected based on preliminary experiments, where several candidate values were tested and the reported configurations showed stable and effective performance across both datasets. Although these settings are empirical, they represent practical tuning commonly adopted in machine learning experiments. We acknowledge that more systematic data-driven or automated parameter selection strategies could further enhance robustness, and this will be an interesting direction for future work.

5.5. Numerical results and analysis

To evaluate model performance more robustly, we perform ten-fold cross-validation with fifty replications for each model. The performance indicators are computed as follows:

$$\overline{P}_k = \frac{1}{10}\sum_{k=1}^{10} P_k, \tag{15}$$

$$\overline{P} = \frac{1}{50}\sum_{k=1}^{50} \overline{P}_k, \tag{16}$$

Table 5. Experimental results for the WDBC dataset (mean ± standard deviation).

Model	AUC (%)	ACC (%)	SP (%)	SE (%)
SVM	97.07 ± 0.004	97.53 ± 0.003	98.99 ± 0.007	**95.15** ± 0.003
Boost-SVM	96.26 ± 0.005	96.52 ± 0.005	97.61 ± 0.005	94.91 ± 0.010
L_1MKL-SVM	97.06 ± 0.003	97.53 ± 0.003	99.08 ± 0.004	95.04 ± 0.006
Ada-L_1MKL-WSVM	96.81 ± 0.006	97.23 ± 0.005	98.53 ± 0.005	95.09 ± 0.010
AUC-Ada-L_1MKL-WSVM	**97.15** ± 0.004	**97.61** ± 0.003	**99.16** ± 0.003	95.13 ± 0.001

Table 6. Experimental results for the WBC dataset (mean ± standard deviation).

Model	AUC (%)	ACC (%)	SP (%)	SE (%)
SVM	96.81 ± 0.002	96.93 ± 0.001	97.28 ± 0.004	**96.35** ± 0.001
Boost-SVM	97.08 ± 0.004	96.72 ± 0.005	95.65 ± 0.008	98.52 ± 0.008
L_1MKL-SVM	97.01 ± 0.003	97.09 ± 0.003	97.35 ± 0.002	96.08 ± 0.006
Ada-L_1MKL-WSVM	95.18 ± 0.006	95.87 ± 0.004	97.59 ± 0.003	92.76 ± 0.113
AUC-Ada-L_1MKL-WSVM	**97.09** ± 0.002	**97.20** ± 0.002	**97.53** ± 0.002	96.64 ± 0.004

where P_k represents the values of AUC, ACC, SP, and SE for the k-th partition, $\overline{P}_k$ denotes the average performance across the ten folds, and $\overline{P}$ is the final performance measure averaged over fifty replications.

The AUC, ACC, SP, and SE of each model in the WDBC and WBC datasets are reported in Tables 5 and 6, respectively. In this section, the third model, denoted as L_1MKL-SVM, is implemented within the L_1-norm multiple kernel learning framework with fixed and equal sample weights, and therefore, serves as the equal sample-weight baseline. In contrast, Ada-L_1MKL-WSVM and AUC-Ada-L_1MKL-WSVM introduce an iterative sample reweighting scheme that increases the weights of misclassified or hard malignant samples, aiming to enhance the sensitivity of the classifier.

The results presented in Tables 5 and 6 show that AUC-Ada-L_1MKL-WSVM consistently achieves the best overall performance in terms of AUC and ACC on both datasets without additional feature extraction. For the WDBC dataset, AUC-Ada-L_1MKL-WSVM attains the highest AUC (97.15%) and ACC (97.61%), while also slightly improving SP and SE compared with L_1MKL-SVM (SP: 99.16% vs. 99.08%; SE: 95.13% vs. 95.04%). For the WBC dataset, AUC-Ada-L_1MKL-WSVM again achieves the highest AUC (97.09%) and ACC (97.20%), and increases SE from 96.08% to 96.64% relative to L_1MKL-SVM, while maintaining a higher SP (97.53% vs. 97.35%).

Inspection of Table 6 further indicates that Boost-SVM yields the highest SE, but suffers from the lowest SP, whereas Ada-L_1MKL-WSVM exhibits the opposite pattern, with the highest SP and the lowest SE. In contrast, AUC-Ada-L_1MKL-WSVM provides a more balanced trade-off between SE and SP, attaining simultaneously high values on both metrics. Overall, AUC-Ada-L_1MKL-WSVM surpasses Ada-L_1MKL-WSVM in most performance aspects across both datasets.

It is worth noting that ensemble learning does not always ensure improved performance. For example, Boost-SVM may underperform standard SVM when the base classifiers are highly correlated, since the lack of diversity reduces the effectiveness of boosting and noisy samples may be overweighted. Similarly, Ada-L_1MKL-WSVM can sometimes perform worse than L_1MKL-SVM, as the mismatch between early sample weight updates and kernel

Table 7. Comparison of existing model results for the WDBC dataset.

Model	ACC (%)
WAUCE (H. F. Wang et al., 2018)	**97.68**
EOA-mK-SVM (Sannasi Chakravarthy et al., 2022)	97.19
eGauss+with PCA (Škrjanc et al., 2022)	95.99
LR-KPCA–LS-SVM (Zhang et al., 2021)	96.00
Fuzzy-ID3+FUZZTDBD (Idris & Ismail, 2021)	94.53
IRFRE (S. Wang et al., 2020)	95.09
AUC-Ada-L_1MKL-WSVM	97.61

Table 8. Comparison of existing model results for the WBC dataset.

Model	ACC (%)
WAUCE (H. F. Wang et al., 2018)	97.10
AK-Boosted C5.0 (Zhang et al., 2021)	95.60
Fuzzy-ID3+FUZZTDBD (Idris & Ismail, 2021)	94.36
IRFRE (S. Wang et al., 2020)	96.44
FA+ELM (Kaya & Kuncan, 2022)	**97.25**
RBFNN (Osman & Aljahdali, 2020)	97.00
AUC-Ada-L_1MKL-WSVM	97.20

optimization may bias the model toward local patterns instead of capturing the global distribution. These observations highlight the importance of designing more robust weighting and kernel adaptation strategies.

In what follows, we compare the performance of AUC-Ada-L_1MKL-WSVM with other models in different literatures for the WDBC and WBC datasets. The raw accuracy metrics of these extant models are systematically presented in Tables 7 and 8, respectively.

Tables 7 and 8 show that the accuracy of our proposed ensemble algorithm is superior to most other models in the literature. For the WDBC dataset, MKL was applied in EOA-mK-SVM (Sannasi Chakravarthy et al., 2022) and ultimately achieved an accuracy of 97.19%.

The experimental results indicate that considering both MKL and the weight of training samples simultaneously is effective. The PCA feature extraction technique was also used in eGuss+with PCA (Škrjanc et al., 2022), and achieved accuracies of 95.99%, respectively.

For the WBC dataset, by comparing the numerical results with the above model, it is shown that our improvement of using SVM as the benchmark model is meaningful. This is because SVM is a very suitable classification model for breast cancer diagnosis. At the same time, we combine the idea of MKL and resampling based on SVM, and improve the resampling method. Overall, our ensemble algorithm provides a more effective method for breast cancer diagnosis.

It should be noted that our method does not always surpass all existing approaches. On the WDBC dataset, the WAUCE method achieves slightly higher accuracy by integrating multiple pre-tuned SVMs and using multi-model voting to reduce variance, which is particularly effective for small-sample data. On the WBC dataset, the FA+ELM method benefits from factor analysis, which reduces noise in low-dimensional features and thus enhances the performance of ELM. By contrast, AUC-Ada-L_1MKL-WSVM focuses on joint optimization of kernels and sample weights, which tends to offer greater advantages in higher-dimensional or more complex scenarios. These observations explain the discrepancies and highlight the specific contexts where our method is most suitable.

6. Conclusions

In this paper, we have developed an AUC-based multi-kernel weighted SVM ensemble algorithm, named AUC-Ada-L_1MKL-WSVM, for breast cancer diagnosis. Our ensemble algorithm not only updates the weights of the training samples, but also optimizes the weights of the combined kernel function. Furthermore, we employ AUC as an optimization criterion to effectively control the weight growth rate of misclassified samples, ensuring improved model stability and generalization. To validate our method, we then applied our ensemble algorithm to the WDBC and WBC datasets, evaluating classification performance across four key metrics: AUC, accuracy, specificity and sensitivity to evaluate our classification performance. Our algorithm achieved outstanding results, with respective performance values of AUC 97.15%, accuracy 97.61%, specificity 99.16%, sensitivity 95.13% on the WDBC and WBC datasets, respectively. These findings demonstrate the efficacy of our proposed ensemble algorithm in breast cancer classification.

In addition, while this work mainly adopts AUC as the guiding metric, we acknowledge that other measures such as MCC could also provide complementary insights. Exploring MCC-based weight updating will be an important direction for our future research.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the National Natural Science Foundation of China [grant numbers 11971302 and 12171307].

ORCID

Guoqiang Wang http://orcid.org/0000-0003-2979-3510

References

Abdar, M., Zomorodi-Moghadam, M., Zhou, X., Gururajan, R., Tao, X., P. D. Barua, & Gururajan, R. (2020). A new nested ensemble technique for automated diagnosis of breast cancer. *Pattern Recognition Letters*, *132*, 123–131. https://doi.org/10.1016/j.patrec.2018.11.004

Akay, M. F. (2009). Support vector machines combined with feature selection for breast cancer diagnosis. *Expert Systems with Applications*, *36*(2), 3240–3247. https://doi.org/10.1016/j.eswa.2008.01.009

Ali, L., Javeed, A., Noor, A., Rauf, H. T., Kadry, S., & Gandomi, A. H. (2024). Parkinson's disease detection based on features refinement through L_1 regularized SVM and deep neural network. *Scientific Reports*, *14*(1), 1333. https://doi.org/10.1038/s41598-024-51600-y

Aljuaid, H., Alturki, N., Alsubaie, N., Cavallaro, L., & Liotta, A. (2022). Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning. *Computer Methods and Programs in Biomedicine*, *223*,106951. https://doi.org/10.1016/j.cmpb.2022.106951

Asri, H., Mousannif, H., Al Moatassime, H., & Noel, T. (2016). Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Procedia Computer Science*, *83*, 1064–1069. https://doi.org/10.1016/j.procs.2016.04.224

Aymaz, S. (2025). Unlocking the power of optimized data balancing ratios: A new frontier in tackling imbalanced datasets. *The Journal of Supercomputing*, *81*(2), 443. https://doi.org/10.1007/s11227-025-06919-2

Bach, F. R., Lanckriet, G. R. G., & Jordan, M. I. (2004). Multiple kernel learning, conic duality, and the SMO algorithm. In *Proceedings of the Twenty-First International Conference on Machine Learning* (p. 6).

Barnett, A. J., Schwartz, F. R., Tao, C. F., Chen, C. F., Ren, Y. H., & Lo, J. Y. (2021). A case-based interpretable deep learning model for classification of mass lesions in digital mammography. *Nature Machine Intelligence*, *3*(12), 1061–1070. https://doi.org/10.1038/s42256-021-00423-x

Beck, A., & Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, *2*(1), 183–202. https://doi.org/10.1137/080716542

Chen, H. L., Yang, B., Liu, J., & Liu, D. Y. (2011). A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. *Expert Systems with Applications*, *38*(7), 9014–9022. https://doi.org/10.1016/j.eswa.2011.01.120

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, *21*(1), 6. https://doi.org/10.1186/s12864-019-6413-7

Gao, S., Lee, C.-H., & Lim, J. H. (2006). An ensemble classifier learning approach to ROC optimization. In *18th International Conference on Pattern Recognition* (pp. 679–682).

Gao, S., & Sun, Q. B. (2007). Improving semantic concept detection through optimizing ranking function. *IEEE Transactions on Multimedia*, *9*(7), 1430–1442. https://doi.org/10.1109/TMM.2007.906597

Ghani, M. U., Alam, T. M., & Jaskani, F. H. (2019). Comparison of classification models for early prediction of breast cancer. In *2019 International Conference on Innovative Computing* (pp. 1–6).

Idris, N. F., & Ismail, M. A. (2021). Breast cancer disease classification using fuzzy-ID3 algorithm with FUZZYDBD method: Automatic fuzzy database definition. *PeerJ Computer Science*, *7*, e427. https://doi.org/10.7717/peerj-cs.427

Jha, C., Li, Y., & Guha, S. (2017). Semiparametric Bayesian analysis of high-dimensional censored outcome data. *Statistical Theory and Related Fields*, *1*(2), 194–204. https://doi.org/10.1080/24754269.2017.1396436

Kapila, R., & Saleti, S. (2023). An efficient ensemble-based machine learning for breast cancer detection. *Biomedical Signal Processing and Control*, *86*,105269. https://doi.org/10.1016/j.bspc.2023.105269

Kashef, R. (2021). A boosted SVM classifier trained by incremental learning and decremental unlearning approach. *Expert Systems with Applications*, *167*,114154. https://doi.org/10.1016/j.eswa.2020.114154

Kaya, Y., & Kuncan, F. (2022). A hybrid model for classification of medical data set based on factor analysis and extreme learning machine: FA+ELM. *Biomedical Signal Processing and Control*, *78*,104023. https://doi.org/10.1016/j.bspc.2022.104023

Le, T. M., & Clarke, B. (2022). Interpreting uninterpretable predictors: Kernel methods, shtarkov solutions, and random forests. *Statistical Theory and Related Fields*, *6*(1), 10–28. https://doi.org/10.1080/24754269.2021.1974157

Levesque, J. C., Durand, A., Gagne, C., & Sabourin, R. (2012). Multi-objective evolutionary optimization for generating ensembles of classifiers in the ROC space. In *Proceedings of the 14th Annual Conference on Genetic and Evolutionary Computation* (pp. 879–886).

Li, J. L., Fine, J. P., & Pencina, M. J. (2017). Multi-category diagnostic accuracy based on logistic regression. *Statistical Theory and Related Fields*, *1*(2), 143–158. https://doi.org/10.1080/24754269.2017.1319105

Li, X. C., Wang, L., & Sung, E. (2008). AdaBoost with SVM-based component classifiers. *Engineering Applications of Artificial Intelligence*, *21*(5), 785–795. https://doi.org/10.1016/j.engappai.2007.07.001

Lin, C. F., & Wang, S. D. (2002). Fuzzy support vector machines. *IEEE Transactions on Neural Networks*, *13*(2), 464–471. https://doi.org/10.1109/72.991432

Liu, N., Qi, E. S., Xu, M., Gao, B., & Liu, G. Q. (2019). A novel intelligent classification model for breast cancer diagnosis. *Information Processing & Management*, *56*(3), 609–623. https://doi.org/10.1016/j.ipm.2018.10.014

Luo, S. H., Dai, Z. A., Chen, T. X., Chen, H. Y., & Jian, L. (2020). A weighted SVM ensemble predictor based on AdaBoost for blast furnace ironmaking process. *Applied Intelligence*, *50*(7), 1997–2008. https://doi.org/10.1007/s10489-020-01662-y

Nanglia, S., Ahmad, M., Khan, F. A., & Jhanjhi, N. (2022). An enhanced predictive heterogeneous ensemble model for breast cancer prediction. *Biomedical Signal Processing and Control*, *72*,103279. https://doi.org/10.1016/j.bspc.2021.103279

Nesterov, Y. (2013). Gradient methods for minimizing composite functions. *Mathematical Programming*, *140*(1), 125–161. https://doi.org/10.1007/s10107-012-0629-5

Osman, A. H., & Aljahdali, H. M. A. (2020). An effective of ensemble boosting learning method for breast cancer virtual screening using neural network model. *IEEE Access*, *8*, 39165–39174. https://doi.org/10.1109/Access.6287639

Polat, K., & Güneş, S. (2007). Breast cancer diagnosis using least square support vector machine. *Digital Signal Processing*, *17*(4), 694–701. https://doi.org/10.1016/j.dsp.2006.10.008

Ragab, D. A., Attallah, O., Sharkas, M., Ren, J., & Marshall, S. (2021). A framework for breast cancer classification using multi-DCNNs. *Computers in Biology and Medicine*, *131*,104245. https://doi.org/10.1016/j.compbiomed.2021.104245

Rakotomamonjy, A., & Grandvalet, Y. (2008). SimpleMKL. *Journal of Machine Learning Research*, *9*(83), 2491–2521.

Ramirez-Morales, A., Salmon-Gamboa, J. U., Li, J., Sanchez-Reyna, A. G., & Palli-Valappil, A. (2023). Boosted support vector machines with genetic selection. *Applied Intelligence*, *53*(5), 4996–5012.

Sahu, B., & Mohanty, S. N. (2021). CMBA-SVM: A clinical approach for Parkinson disease diagnosis. *International Journal of Information Technology*, *13*(2), 647–655. https://doi.org/10.1007/s41870-020-00569-8

Sannasi Chakravarthy, S. R., Rajaguru, H., & Chidambaram, S. (2022). Processing of Wisconsin breast cancer data using Ebola optimization algorithm with mixture kernel SVM. In *2022 Smart Technologies, Communication and Robotics* (pp. 1–4).

Sharma, A., Kaur, S., Memon, N., Fathima, A. J., Ray, S., & Bhatt, M. W. (2021). Alzheimer's patients detection using support vector machine (SVM) with quantitative analysis. *Neuroscience Informatics*, *1*(3),100012. https://doi.org/10.1016/j.neuri.2021.100012

Škrjanc, I., Andonovski, G., Iglesias, J. A., Sanchis, A., & Lughofer, E. (2022). Evolving Gaussian on-line clustering in social network analysis. *Expert Systems with Applications*, *207*,117881. https://doi.org/10.1016/j.eswa.2022.117881

Wang, F., Li, Z. H., He, F., Wang, R., Yu, W. Z., & Nie, F. P. (2019). Feature learning viewpoint of AdaBoost and a new algorithm. *IEEE Access*, *7*,149890–149899. https://doi.org/10.1109/ACCESS.2019.2947359

Wang, H. F., Zheng, B. C., Yoon, S. W., & Ko, H. S. (2018). A support vector machine-based ensemble algorithm for breast cancer diagnosis. *European Journal of Operational Research*, *267*(2), 687–699. https://doi.org/10.1016/j.ejor.2017.12.001

Wang, J. J., He, F., & Sun, S. H. (2023). Construction of a new smooth support vector machine model and its application in heart disease diagnosis. *PloS ONE*, *18*(2), e0280804. https://doi.org/10.1371/journal.pone.0280804

Wang, S., Wang, Y., Wang, D., Yin, Y., Wang, Y., & Jin, Y. (2020). An improved random forest-based rule extraction method for breast cancer diagnosis. *Applied Soft Computing*, *86*, 105941. https://doi.org/10.1016/j.asoc.2019.105941

Xie, X. J., Luo, K. Y., & Wang, G. Q. (2022). A new L_1 multi-kernel learning support vector regression ensemble algorithm with AdaBoost. *IEEE Access*, *10*, 20375–20384. https://doi.org/10.1109/ACCESS.2022.3151672

Yang, X., Song, Q., & Cao, A. (2005). Weighted support vector machine for data classification. In *Proceedings 2005 IEEE International Joint Conference on Neural Networks* (Vol. 2, pp. 859–864).

Zhang, J., Chen, L., Tian, J., Abid, F., Yang, W., & Tang, X. (2021). Breast cancer diagnosis using cluster-based undersampling and boosted C5.0 algorithm. *International Journal of Control, Automation and Systems*, *19*(5), 1998–2008. https://doi.org/10.1007/s12555-019-1061-x

Zhang, X. L., & Ren, F. (2008). Improving SVM learning accuracy with AdaBoost. In *2008 Fourth International Conference on Natural Computation* (pp. 221–225).

Zheng, B. C., Yoon, S. W., & Lam, S. S. (2014). Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms. *Expert Systems with Applications*, *41*(4), 1476–1482. https://doi.org/10.1016/j.eswa.2013.08.044