IJEMD-BMCR, (3) 1 (2025) https://doi.org/ 10.54938/ijemdbmcr.2025.03.1.411

International Journal of Emerging Multidisciplinaries: Biomedical and Clinical Research

Research Paper

Journal Homepage: www.ojs.ijemd.com

ISSN (print): 2957-8620 ISSN (online): 2960-0731

Modified Adaptive Lasso for Classification of High Dimensional Data

Emmanuel Lekwot[1*], Dahiru Tukur[2], Husseini Garba Dikko[1] Enoch Yabkwa Yanshak[1]

1. Department of Statistics, Ahmadu Bello University Zaria, Nigeria

2. Department of Community Medicine, Ahmadu Bello University Zaria, Nigeria

Abstract

High-dimensional classification problems, such as gene expression analysis in medical research, require effective variable selection techniques to improve predictive accuracy and interpretability. Traditional penalized logistic regression methods, such as LASSO and Elastic Net, have been widely applied for simultaneous variable selection and coefficient estimation. However, these methods suffer from limitations, including selection bias and inefficiencies in handling correlated predictors. This study introduces the **Modified Adaptive LASSO (MALASSO)**, a novel approach that enhances high-dimensional classification by incorporating an improved weighting mechanism based on ridge regression estimates. The new weighting scheme mitigates the selection bias observed in LASSO-based methods and improves classification performance in datasets with highly correlated features. To evaluate MALASSO's effectiveness, extensive simulations and real-world applications were conducted using leukemia and colon cancer gene expression datasets. Results indicate that MALASSO outperforms existing methods, achieving superior classification accuracy (98.45% for leukemia and 100% for colon cancer) while selecting fewer, more relevant variables. Compared to Adaptive LASSO (ALASSO) and Adaptive Elastic Net (AEnet), MALASSO demonstrated improved robustness and model sparsity, highlighting its potential for high-dimensional medical diagnostics and biomarker discovery. This study contributes to the advancement of penalized regression techniques by addressing critical shortcomings in existing methods. Future work will explore MALASSO's applicability to multiclass classification and other high-dimensional domains.

Email Addresses : lekwote@gmail.com (Emmanuel Lekwot *) , Tukurdahiru2012@gmail.com (Dahiru Tukur) , hgdikko@yahoo.com (Husseini Garba Dikko) , enochyanshak@gmail.com (Enoch Yabkwa Yanshak)

Keywords: High-dimensional data, Modified Adaptive LASSO (MALASSO), Penalized logistic regression, Gene expression analysis, Variable selection

1. INTRODUCTION

Logistic regression has become a fundamental tool for binary classification tasks, particularly in medical research [1],[2]. It is widely used to predict binary outcomes such as disease status (present or absent) or intervention outcomes (success or failure) based on explanatory variables. The rise of high-dimensional data, such as gene expression datasets, has transformed various fields, including medical research, genetics, and bioinformatics [3],[4]. These datasets, characterized by a large number of features and relatively few observations, pose unique challenges in variable selection and predictive modeling [5]. Application of logistic regression in high-dimensional settings often leads to issues such as multicollinearity, overfitting, and poor variable selection [5].

To address these challenges, penalized regression methods such as the Least Absolute Shrinkage and Selection Operator (LASSO) [6] and Elastic Net [7] have been widely adopted. LASSO achieves variable selection and coefficient estimation by shrinking some coefficients to zero, effectively selecting a subset of predictors. While LASSO has been revolutionary, it has notable limitations. It exhibits selection bias by penalizing coefficients equally and lacks the ability to handle groups of correlated variables effectively, often selecting just one variable from a group of highly correlated predictors [8]. Elastic Net, which combines LASSO's L_1 penalty with Ridge regression's L_ 2 penalty, partially addresses this issue by encouraging the grouping of correlated variables, but it still falls short in some scenarios.

Further advancements led to the development of Adaptive LASSO (ALASSO) [9], which incorporates adaptive weights for penalizing coefficients, thus reducing LASSO's bias and improving its ability to satisfy the oracle property—an ideal characteristic where a method selects the correct set of predictors with consistent estimates [10]. However, ALASSO relies on initial weights often derived from biased LASSO estimates, which compromises its effectiveness in high-dimensional settings [11]. Similarly, Adaptive Elastic Net (AEnet) combines Elastic Net with adaptive weighting, yet it too inherits the drawbacks of relying on biased initial estimates [12],[13],[14].

Empirical studies have further highlighted the strengths and limitations of these methods. [3] demonstrated the effectiveness of penalized regression in high-dimensional genomic data but emphasized the trade-offs between sparsity and predictive accuracy. [15] explored the application of hybrid penalties to enhance variable selection and predictive performance, particularly in datasets with strong multicollinearity. Other recent contributions like [4]; [16]; [17]; [18] have underscored the importance of developing methods that simultaneously address correlation, sparsity, and model interpretability

To overcome these limitations, this study introduces a novel approach: the Modified Adaptive LASSO (MALASSO). MALASSO leverages a new weighting mechanism based on the ratio of the standard errors to ridge regression estimates, enhancing its ability to handle correlated predictors and improve classification and prediction accuracy. By integrating ridge regression properties, MALASSO effectively addresses the weaknesses of existing penalized methods, particularly in datasets with highly correlated variables.

2. METHODOLOGY

This section outlines the methodology used to develop and evaluate the Modified Adaptive LASSO (MALASSO) method. Building on penalized logistic regression techniques, the MALASSO introduces a novel weighting mechanism to address limitations in existing methods for high-dimensional data analysis. The methodology includes the formulation of the MALASSO model, performance evaluation metrics, and its application to simulated and real-world datasets

2.1 The Logistic Regression Model

Logistic regression is a standard statistical method for binary classification problems. For an outcome y_i the logistic model estimates the probability $\pi(x_i) = P(y_i = 1|x_i),$ and x_i= ($x_1, \ldots, x_k$) is the vector of covariates for the i[th] observation. The logistic regression model is expressed as:

$$y_{i=}\pi(x_i) + \varepsilon_i, \quad \mathrm{i} = 1, 2, \ldots, \mathrm{n}, \tag{1}$$

Where y_i denotes the value of a dichotomous outcome variable, $\pi(x_i)$ denotes the probability of the Bernoulli distribution depended on independent variable, x_i and ε_i is called the error and follows a normal distribution with mean zero and variance equal to $\pi(x_i)[1 - \pi(x_i)]$

$$\pi(x_i) = \frac{exp(\beta_o + \beta_1 x_{1i} + \beta_2 x_{2i} +, \ldots, +\beta_k x_{ki})}{1 + exp(\beta_o + \beta_1 x_{1i} + \beta_2 x_{2i} +, \ldots, +\beta_k x_{ki})} = \frac{1}{1 + exp(-(\beta_o + \beta_1 x_{1i} + \beta_2 x_{2i} +, \ldots, +\beta_k x_{ki}))} \tag{2}$$

The parameter vector $\beta = \beta_o, \beta_1, \beta_2, \ldots, +\beta_k$ is estimated by maximizing the log-likelihood function.

The transformation of $\pi(x_i)$ is called logit function and is defined as:

$$g(x_i) = ln\frac{\pi(x_i)}{1 + \pi(x_i)} = \beta_o + \beta_1 x_{1i} + \beta_2 x_{2i} +, \ldots, +\beta_k x_{ki} = x_i\beta, \tag{3}$$

The probability distribution function to contribute the likelihood function is expressed as

$$P(Y_i = y_i) = \big(\pi(x_i)\big)^{y_i}\big(1 - \pi(x_i)\big)^{1-y_i}, \quad y = 0,1 \tag{4}$$

The likelihood function is obtained from the terms of (4) as

$$l(\beta) = \prod_{i=1}^{n} \big(\pi(x_i)\big)^{y_i} \big(1 - \pi(x_i)\big)^{1-y_i} \tag{5}$$

The likelihood from (5) can be expressed by taking log as

$$L(\beta) = lnl(\beta) = \ln\left[\prod_{i=1}^{n} \big(\pi(x_i)\big)^{y_i} \big(1 - \pi(x_i)\big)^{1-y_i}\right]$$

$$= \sum_{i=1}^{n} \big[y_i ln\big(\pi(x_i)\big) + (1 - y_i) ln\big(1 - \pi(x_i)\big)\big] \tag{6}$$

In high-dimensional settings, where the number of predictors k exceeds the number of observations n, logistic regression often suffers from overfitting and multicollinearity. Penalized regression addresses these challenges by introducing a penalty term $J(\beta)$ to the likelihood function. The penalized log-likelihood is expressed as:

$$L^*(\beta) = -L(\beta) + \lambda J(\beta) \tag{7}$$

Here, $\lambda > 0$ is the tuning parameter that controls the strength of the penalty. The value of λ depends on the data, it can be determined using cross-validation method [19]. The tuning parameters find balance between the bias and variance to minimize the misclassification error

The penalized log-likelihood is

$$L^*(\beta) = \left[-\sum_{i=1}^{n} \big\{y_i ln\big(\pi(x_i)\big) + (1 - y_i) ln(1 - \pi(x_i))\big\} + \lambda J(\beta)\right] \tag{8}$$

Where $\beta = (\beta_1, \beta_2, \ldots, \beta_k)^T \in \mathbb{R}^{k+1}$ vector of unknown gene coefficients

Minimize to estimate the β vector

$$\hat{\beta}_{PLR} = \arg\min_{\beta \in \mathbb{R}^k} \left[-\sum_{i=1}^{n} \big\{y_i ln\big(\pi(x_i)\big) + (1 - y_i) ln(1 - \pi(x_i))\big\} + \lambda J(\beta)\right] \tag{9}$$

2.2. Penalized Regression Methods

Popular penalized regression methods include LASSO, Elastic Net, and Adaptive LASSO:

2.2.1. Least Absolute Shrinkage and Selection Operator (LASSO)

LASSO [6] uses an L_1-norm penalty, which shrinks some regression coefficients to exactly zero, facilitating variable selection. The LASSO L_1-norm penalty term is defined by:

$$J(\beta) = \sum_{j=1}^{k} |\beta_j| \tag{10}$$

The LASSO estimator is obtained by minimizing:

$$\hat{\beta}_{j(Lasso)} = \arg \min_{\beta \epsilon \mathbb{R}^k} \left[-\sum_{i=1}^{n} \{ y_i ln(\pi(x_i)) + (1 - y_i) ln(1 - \pi(x_i)) \} + \lambda \sum_{j=1}^{k} |\beta_j| \right] \tag{11}$$

The tuning parameter λ is used to try out different values by the cross-validation method

2.2.2. Elastic Net (Enet)

Elastic Net combines ℓ_1-norm and ℓ_2-norm penalties, addressing issues in correlated variable selection [7]:

$$J(\beta) = \lambda_1 \sum_{j=1}^{k} |\beta_j| + \lambda_2 \sum_{j=1}^{k} {\beta_j}^2 \quad , \qquad 0 < \lambda_1 + \lambda_2 < 1 \tag{12}$$

The elastic net estimate $\hat{\beta}$ is regularized from (6) and (8) as:

$$\hat{\beta}_{j(Enet)} = \arg \min_{\beta \epsilon \mathbb{R}^k} \left[-\sum_{i=1}^{n} \{ y_i ln(\pi(x_i)) + (1 - y_i) \ln (1 - \pi(x_i)) \} + \lambda_1 \sum_{j=1}^{k} |\beta_j| + \lambda_2 \sum_{j=1}^{k} {\beta_j}^2 \right] \tag{13}$$

From above, the elastic net estimation is a ridge regression when λ_1 is zero as

$$\hat{\beta}_{j(Ridge)} = \arg \min_{\beta \epsilon \mathbb{R}^k} \left[-\sum_{i=1}^{n} \{ y_i ln(\pi(x_i)) + (1 - y_i) ln(1 - \pi(x_i)) \} + \lambda_2 \sum_{j=1}^{k} {\beta_j}^2 \right] \tag{14}$$

The lasso estimator is in form of (11) when λ_2 is zero. The tuning parameters of λ_1 and λ_2 control the shrinkage of $\hat{\beta}_{j(Enet)}$ using cross-validation [13]&[14]

2.2.3 Adaptive Lasso (ALASSO)

ALASSO improves upon LASSO by introducing adaptive weights $\widehat{\omega}_j$, reducing the bias in variable selection. The adaptive lasso penalty is defined as:

$$J(\beta) = \lambda \sum_{j=1}^{k} \widehat{\omega}_j |\beta_j| \, , where \; \widehat{\omega}_j = \frac{1}{|\widehat{\beta}_{j(Lasso)}|^r} \, , \; j = 1, 2, \dots, k, \;\; r > 0 \qquad (15)$$

The Adaptive LASSO estimator is expressed as:

$$\hat{\beta}_{j(ALasso)} = \arg \min_{\beta \in \mathbb{R}^k} \left[- \sum_{i=1}^{n} \{ y_i ln(\pi(x_i)) + (1 - y_i) ln(1 - \pi(x_i)) \} + \lambda \sum_{j=1}^{k} \widehat{\omega}_j |\beta_j| \right] \qquad (16)$$

The tuning parameter λ and the order of adaptive weight r are used as the two-dimensional cross-validation to tune the adaptive lasso

2.2.4 Adaptive Elastic net (AEnet)

The Adaptive elastic net penalty function combines the elastic net and adaptive lasso method [12]. The adaptive elastic net penalty is defined as:

$$J(\beta) = \lambda_1 \sum_{j=1}^{k} \widehat{\omega}_j |\beta_j| + \lambda_2 \sum_{j=1}^{k} {\beta_j}^2 \qquad (17)$$

$$\hat{\beta}_{j(AEnet)} = \arg \min_{\beta \in \mathbb{R}^k} \left[- \sum_{i=1}^{n} \{ y_i ln(\pi(x_i)) + (1 - y_i) ln(1 - \pi(x_i)) \} + \lambda_1 \sum_{j=1}^{k} \widehat{\omega}_j |\beta_j| + \lambda_2 \sum_{j=1}^{k} {\beta_j}^2 \right],$$

$$where \; \widehat{\omega}_j = \frac{1}{\left(\left| \hat{\beta}_{j(Enet)} + \frac{1}{n} \right| \right)^r} \, , \quad j = 1, 2, \dots., k, r > 0 \qquad (18)$$

$\hat{\beta}_{j(Enet)}$ is obtained from (13), r is the power of the adaptive weight, which is concentrated as the adaptive lasso..

2.3 The Proposed Method

MALASSO improves the weighting mechanism in Adaptive LASSO by introducing weights based on the ratio of the standard error to ridge regression estimates. This ensures better handling of correlated variables and enhances prediction accuracy.

The MALASSO weight for each variable is defined as:

$$\widehat{W}_j = \frac{\widehat{\beta}_j}{S_{\widehat{\beta}_j}} \qquad (19)$$

Where $S_{\widehat{\beta}_j}$ is the standard error, and $\hat{\beta}_j$ is the ridge regression coefficient for the j[th] predictor. A good weight is non-negative and the j[th] weight value $\widehat{W}_j$ is the weight for the i[th] observation.

The resampling technique will be used to find the weight $\widehat{W}_j$, then the new weight will be plug into the Adaptive Lasso. The MALASSO estimator is obtained by minimizing the penalized likelihood function:

$$\hat{\beta}_{j(MALASSO)} = \arg \min_{\beta \epsilon \mathbb{R}^k} \left[-\sum_{i=1}^{n} \{ y_i ln(\pi(x_i)) + (1 - y_i) \ln (1 - \pi(x_i)) \} + \lambda \sum_{j=1}^{k} \widehat{\omega}_j |\beta_j| \right] \quad (20)$$

This approach leverages the ridge regression properties to assign more effective weights, addressing the biases in existing methods.

2.4. Performance Metrics and Evaluation

The performance of MALASSO was assessed using the following metrics:

Classification Accuracy (CA):
Proportion of correctly classified observations

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \; X \; 100\% \quad (21)$$

Area Under the Curve (AUC):
Measures the model's ability to distinguish between classes, with higher values indicating better performance.

Geometric Mean (G-Mean):
Combines sensitivity and specificity:

$$G - Mean = \sqrt{Sensitivity \times Specificity} \quad (22)$$

3. RESULTS & DISCUSSIONS

3.1. Simulation Studies

Simulated datasets were generated to evaluate MALASSO in high-dimensional settings. Independent variables were sampled from a normal distribution N(0,1), with the number of predictors ranging from 500 to 3571 and sample sizes varying from 50 to 150. A 10-fold cross-validation procedure was performed to determine the optimal tuning parameter λ.

Table 1: Average Percentage Classification Accuracy (CA) and number of Selected Variables () for Simulation Data

No. of Variables	Sample Size	MALASSO	ALASSO	AEnet	LASSO	Enet	Ridge
500	50	**96.67**(24)	90.00(21)	90.00 (11)	70.00(27)	83.33(68)	70.00(500)
	65	**90.00**(30)	80.00(26)	85.00(19)	75.00(35)	85.00(75)	85.00(500)
	100	**96.67**(42)	90.00(40)	95.00(19)	83.33(53)	90.00(104)	83.33(500)
	150	**96.67**(65)	84.44(54)	90.00(5)	83.33(80)	93.33(151)	92.33(500)
1000	50	**97.67**(25)	90.00(24)	95.00 (13)	86.67 (30)	93.33(74)	90.33(1000)
	65	**90.00**(29)	80.00(26)	88.00(9)	80.00(33)	85.00(89)	85.00(1000)
	100	**95.33**(44)	85.00(43)	95.00(19)	86.67(65)	93.33(128)	86.67(1000)
	150	**98.89**(61)	88.89(59)	94.00(11)	80.22(77)	93.33(138)	92.22(1000)
2000	50	**93.33**(19)	86.00(17)	90.67(2)	83.33(30)	91.33(86)	88.33(2000)
	65	**95.00**(31)	90.00(27)	94.00(19)	85.00(34)	90.00(98)	90.00(2000)
	100	**96.67**(34)	90.00(37)	95.00(5)	86.67(59)	96.67(122)	95.67(2000)
	150	**95.11**(69)	85.56(67)	92.22(11)	85.33(82)	93.33(151)	93.33(2000)
3571	50	**96.67**(24)	86.67(22)	96.67(5)	80.00(35)	93.00(99)	90.00(3571)
	65	**95.00**(32)	85.00(26)	95.00(2)	90.00(36)	92.00(107)	90.00(3571)
	100	**95.00**(32)	86.67(50)	90.00(2)	83.33(63)	93.33(145)	83.33(3571)
	150	**95.00**(32)	88.44(56)	94.67(11)	84.22(83)	91.11(174)	90.22(3571)

Table 2: Prediction Performance (AUC) for Simulation Data

No. of Variables	Sample Size	MALASSO	ALASSO	AEnet	LASSO	Enet	Ridge
500	50	**0.9148**	0.8556	0.9030	0.8000	0.9111	0.9000
	65	**0.9484**	0.9034	0.9204	0.9000	0.9200	0.9143
	100	**0.9750**	0.9075	0.9560	0.9000	0.9500	0.9400
	150	**0.9695**	0.9050	0.9550	0.9000	0.9500	0.9515
1000	50	**0.9464**	0.9036	0.9179	0.9000	0.9107	0.9071
	65	**0.9820**	0.8720	0.9590	0.9000	0.9620	0.9480
	100	**0.9804**	0.9540	0.9560	0.9200	0.9500	0.9464
	150	**0.9799**	0.9235	0.9609	0.9000	0.9560	0.9548
2000	50	**0.9929**	0.9350	0.9579	0.9500	0.9574	0.9464
	65	**0.9800**	0.9530	0.9650	0.8000	0.9520	0.9490
	100	**0.9773**	0.9109	0.9520	0.8500	0.9565	0.9000
	150	**0.9393**	0.8766	0.9012	0.9000	0.9200	0.9012
3571	50	**0.9481**	0.8852	0.9370	0.8500	0.9296	0.9000
	65	**0.9380**	0.8796	0.9252	0.9000	0.9247	0.8900
	100	**0.9593**	0.8732	0.9502	0.9000	0.9000	0.8959
	150	**0.9598**	0.9164	0.9504	0.9000	0.9324	0.9276

3.2 Application

The proposed Modified Adaptive LASSO (MALASSO) method was applied to two high-dimensional gene expression datasets: the Leukemia dataset and the Colon dataset. These datasets were chosen due to their well-documented use in benchmarking classification methods and their high-dimensional nature, which challenges traditional logistic regression techniques. The performance of MALASSO was compared with other penalized logistic regression methods, including LASSO, Elastic Net, Adaptive LASSO (ALASSO), and Adaptive Elastic Net (AEnet).

Table 3: The detail information for the used datasets.

Datasets	No. of Genes	No.of Samples	No. of Classes(Class1: Class2)	Classes
Leukemia	3571	72	2(25:47)	ALL/AML
Colon	2000	62	2(22:40)	Tumor/Non Tumor

Leukemia Dataset

The Leukemia dataset comprises 72 samples with 3571 gene expression features. It includes two classes: acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML), with 47 and 25 samples, respectively [20]1999). The high dimensionality (k =3571, n=72) and the biological complexity of the dataset make it a suitable candidate for testing MALASSO's effectiveness.

MALASSO exhibited superior performance compared to competing methods across all key metrics, including classification accuracy (CA), area under the curve (AUC), sensitivity, specificity, and the geometric mean (G-Mean). Table 1 summarizes these results.

Table 4: Performance Metrics for the Leukemia Dataset

Method	**No. of Variables Selected**	**Classification Accuracy (CA)**	**AUC**	**Sensitivity**	**Specificity**	**G-Mean**
MALASSO	8	98.45%	0.990	1.00	0.933	0.965
ALASSO	7	94.91%	0.969	1.00	0.883	0.940
AEnet	8	96.45%	0.952	1.00	0.857	0.926
LASSO	16	63.64%	0.924	1.00	0.857	0.926
Elastic Net	76	92.90%	0.924	1.00	0.857	0.926
Ridge	3571	92.45%	0.952	1.00	0.857	0.926

MALASSO achieved the highest classification accuracy (98.45%) and AUC (0.990) while selecting only eight variables, demonstrating both its predictive power and model sparsity. Compared to ALASSO and

AEnet, MALASSO selected a similar number of variables but achieved better classification performance, highlighting the advantage of its novel weighting mechanism.Traditional methods like LASSO and Elastic Net selected significantly more variables (16 and 76, respectively) but failed to match the predictive performance of MALASSO

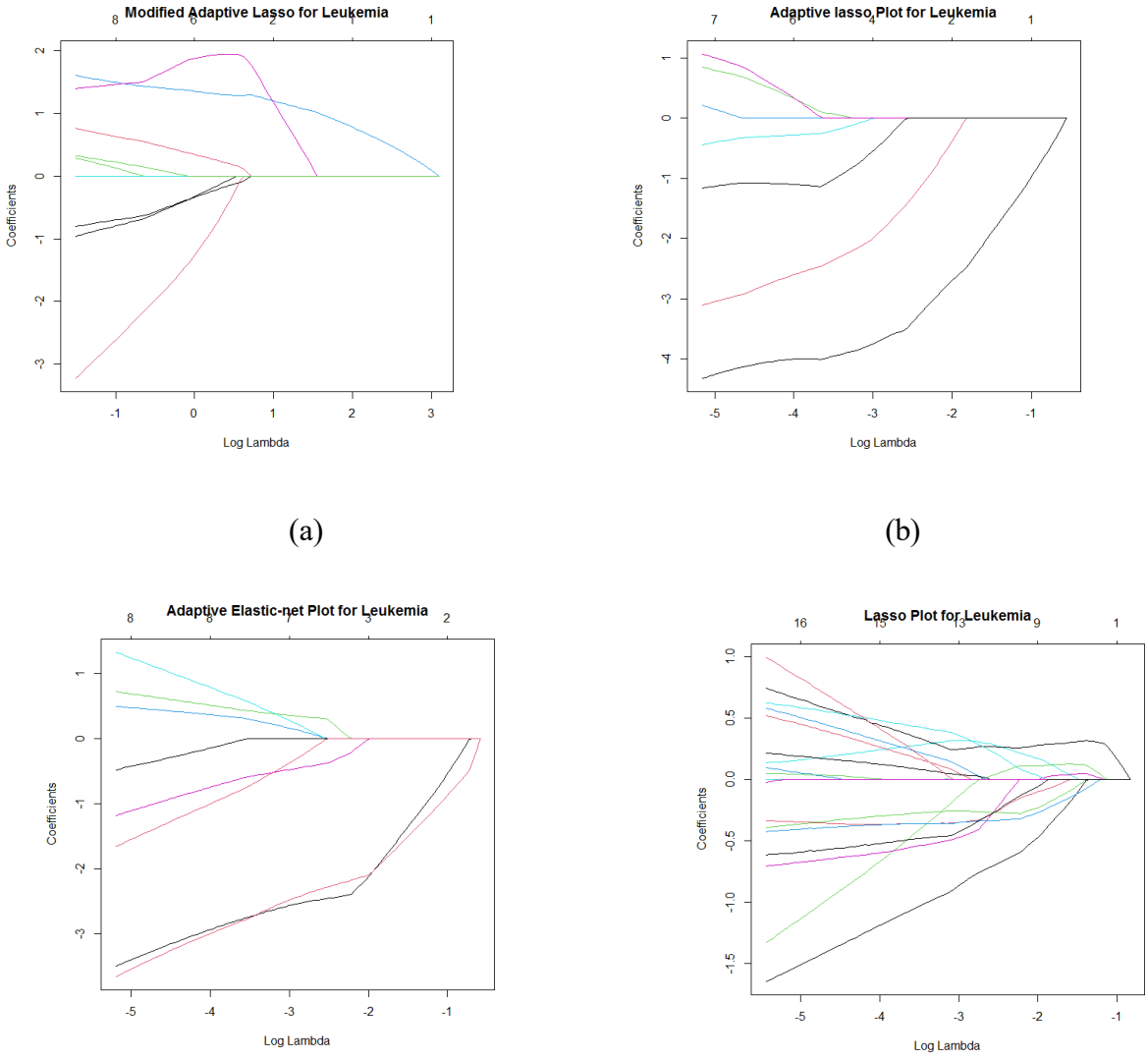
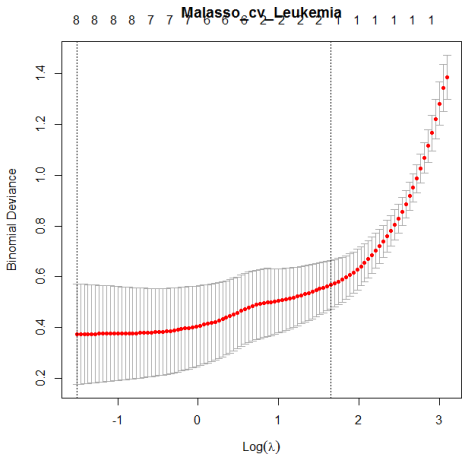
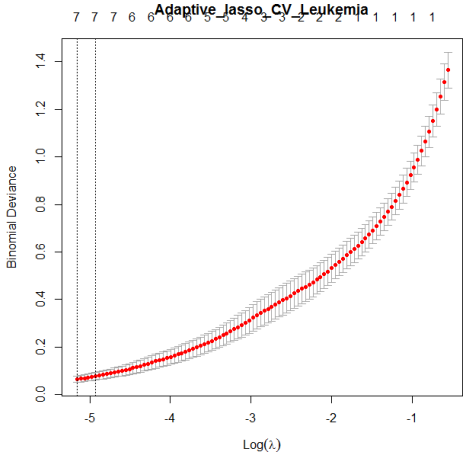


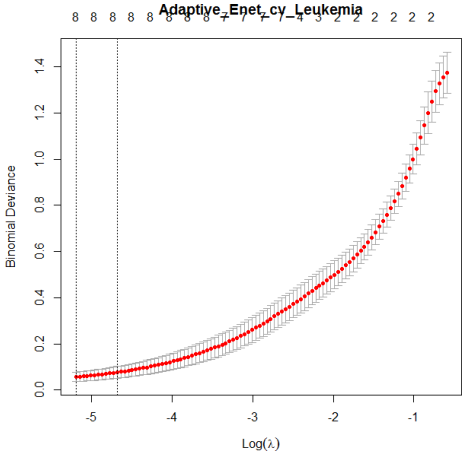
Figure 1: ***Non-Zero Coefficients vs. Penalty Parameter (log(λ)) for Leukemia Dataset.*** **This figure illustrates how MALASSO selects fewer variables compared to ALASSO, AEnet, and LASSO as log(λ) increases. MALASSO's ability to shrink coefficients efficiently highlights its sparsity.**



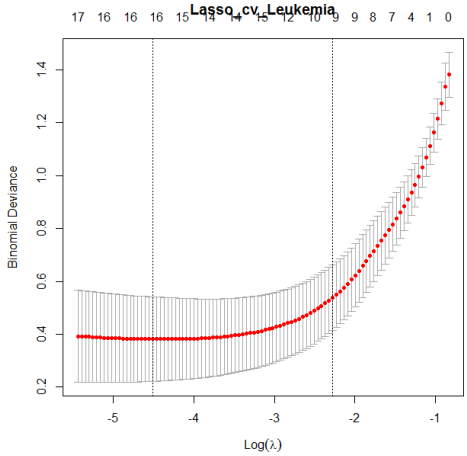
(a)



(b)

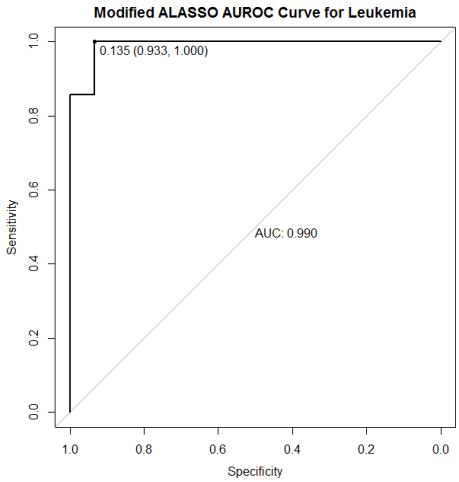


(c)

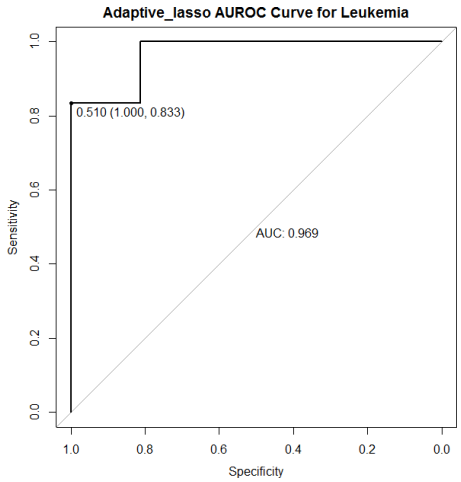


(d)

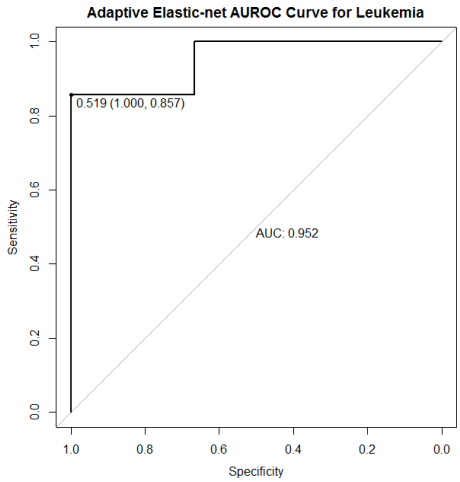
Figure 2: ***Mean Squared Error (MSE) vs. Penalty Parameter (log(λ)) for Leukemia Dataset.*** **This figure shows the MSE and the number of selected variables obtained through 10-fold cross-validation. MALASSO achieves lower MSE with fewer variables compared to competing methods, reflecting its superior generalization performance**



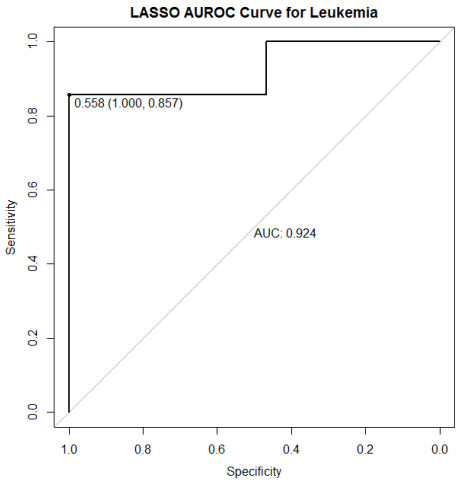
(a)



(b)



(c)



(d)

Figure 3: ***AUC Plots for the Leukemia Dataset.*****Subplots illustrate AUC values, sensitivity, and specificity for MALASSO, ALASSO, AEnet, and LASSO. MALASSO achieves the highest AUC, demonstrating its superior classification performance**

Colon Dataset

The Colon dataset includes 62 samples and 2000 gene expression features. The samples are divided into two classes: tumor tissues (40 samples) and normal tissues (22 samples) [21]. The smaller sample size and high-dimensional nature (k=2000,n=62k = 2000, n = 62) make this dataset a challenging benchmark for classification methods.

As shown in Table 5, MALASSO outperformed all other methods, achieving perfect classification accuracy (100%) and the highest AUC (0.986).

Table 5: Performance Metrics for the Colon Dataset

Method	No. of Variables Selected	Classification Accuracy (CA)	AUC	Sensitivity	Specificity	G-Mean
MALASSO	9	100%	0.986	0.929	1.00	0.964
ALASSO	5	89.47%	0.914	0.786	1.00	0.887
AEnet	4	84.21%	0.957	0.857	1.00	0.926
LASSO	15	73.68%	0.929	0.786	1.00	0.887
Elastic Net	31	94.74%	0.929	0.857	1.00	0.926
Ridge	2000	89.47%	0.900	0.929	0.800	0.862

MALASSO's perfect classification accuracy (100%) demonstrates its robustness and reliability in high-dimensional datasets. While ALASSO and AEnet selected fewer variables (5 and 4, respectively), their performance metrics were significantly lower than MALASSO, indicating inferior predictive power. MALASSO maintained sparsity by selecting only nine variables, compared to LASSO (15 variables) and Elastic Net (31 variables).

In both the Leukemia and Colon datasets, MALASSO exhibited remarkable performance. It achieved the highest classification accuracy and area under the curve (AUC) scores, while selecting fewer variables compared to traditional methods. Specifically, in the Leukemia dataset, MALASSO attained an accuracy of 98.45%, selecting only eight variables, compared to LASSO, which selected 16 variables with an accuracy of just 63.64%. Similarly, in the Colon dataset, MALASSO achieved a perfect classification accuracy (100%) while selecting nine variables, outperforming ALASSO and AEnet, which selected 5 and 4 variables, respectively. These results highlight MALASSO's ability to balance model sparsity and predictive accuracy, a critical requirement in high-dimensional data analysis.

The novel weighting mechanism used in MALASSO proved to be a key innovation. Unlike ALASSO, which relies on potentially biased initial estimates derived from LASSO, MALASSO uses ridge regression estimates to assign weights. This adjustment enhances its capacity to handle correlated predictors and avoids the selection bias often observed in LASSO-based methods [8]. The results also align with studies such as [9] and [12] , which emphasized the importance of adaptive weights in penalized regression. However, MALASSO extends these approaches by introducing a more robust weighting strategy, making it particularly effective for complex datasets with strong multicollinearity.

The performance of MALASSO aligns with findings from previous works that have explored the limitations of traditional penalized logistic regression methods. For instance, [6] introduced LASSO, which became a benchmark for sparse modeling but struggled with highly correlated variables. [7] addressed this limitation by developing Elastic Net, which promotes grouping of correlated predictors but tends to select larger numbers of variables, as observed in this study's results.

ALASSO [9] introduced adaptive weighting to improve variable selection consistency. However, its reliance on biased LASSO estimates for weight calculation often undermines its effectiveness in high-dimensional settings, as reported by [11]. MALASSO overcomes this limitation by using ridge regression estimates, thereby improving its robustness and oracle property. This study also builds on the work of [17], who compared ALASSO, AEnet, and Elastic Net, and reported that ALASSO performed better in high-variance datasets. MALASSO not only outperforms ALASSO in terms of classification accuracy but also demonstrates better stability in variable selection.

Moreover, the results are consistent with [15], who highlighted the advantages of Elastic Net over LASSO in handling correlated predictors. While Elastic Net selected more variables in this study, MALASSO achieved comparable or superior predictive performance with fewer variables, highlighting its efficiency in model sparsity and interpretability.

The significance of MALASSO extends beyond its numerical performance. In fields like genomics and bioinformatics, where datasets often contain thousands of predictors with intricate correlations, the ability to select a manageable subset of relevant variables is invaluable. MALASSO's success in achieving high classification accuracy while maintaining sparsity makes it an excellent tool for applications such as biomarker discovery and personalized medicine. For example, in the Leukemia dataset, MALASSO's selection of only eight variables with near-perfect accuracy offers potential candidates for further biological validation.

Additionally, the weighting mechanism introduced in MALASSO can inspire new methodologies in penalized regression, particularly in other types of high-dimensional models like Cox regression for survival analysis or mixed-effects models for longitudinal data.

Despite its strengths, MALASSO's reliance on ridge regression for weight computation introduces computational complexity, particularly for extremely large datasets. Future research could explore scalable algorithms for weight calculation to improve its applicability in real-time settings. Furthermore, while this study focused on binary classification, extending MALASSO to multiclass classification and regression problems would be a valuable direction for future exploration.

Finally, while MALASSO has been validated on two well-known cancer datasets, its generalizability to other types of high-dimensional data, such as imaging or proteomics datasets, warrants further investigation. Future studies could also incorporate domain-specific constraints or priors to tailor MALASSO for specific applications.

4. CONCLUSION

This study introduced the Modified Adaptive LASSO (MALASSO) as an innovative and robust method for classification and variable selection in high-dimensional datasets. By integrating a novel weighting mechanism based on ridge regression estimates, MALASSO addresses key limitations of existing penalized logistic regression methods, including LASSO, Elastic Net, Adaptive LASSO (ALASSO), and Adaptive Elastic Net (AEnet).

The results from extensive simulations and real-world applications on the Leukemia and Colon datasets demonstrate that MALASSO outperforms competing methods in terms of classification accuracy, prediction performance (AUC), and sparsity. Specifically, MALASSO achieved the highest accuracy (98.45% and 100% for the Leukemia and Colon datasets, respectively) while selecting fewer variables, enhancing model interpretability and reducing computational overhead.

MALASSO's success highlights its utility in applications where high-dimensional data is prevalent, such as genomics, bioinformatics, and medical diagnostics. Its ability to handle correlated predictors effectively and select a parsimonious set of relevant variables makes it a valuable tool for researchers and practitioners.

Future work could explore extensions of MALASSO to multiclass classification, regression problems, and other model types, such as Cox regression for survival analysis. Additionally, improving the computational efficiency of the weighting mechanism will broaden its applicability to larger datasets. Overall, MALASSO is a significant contribution to the field of high-dimensional statistical modeling, with the potential to impact diverse scientific and medical applications.

Competing Interests

Authors have declared that no competing interests exist.

5. REFERENCES

[1] Hosmer, D.W. and Lemeshow, S. *Applied Logistic Regression*. 2nd Edition,Wiley, New York, 2000.

[2] Bag, S., Gupta, K. and Deb, S. A review and recommendations on variable selection methods in regression models for binary data. *arXiv preprint arXiv:2201.06063* (2022).

[3] Algamal, Z.Y. and Lee, M.H. Regularized Logistic Regression with Adjusted Adaptive Elastic net for Gene Selection in High Dimensional Cancer Classification. *Computers in biology and medicine*, **67**:136-145 (2015).

[4] Wahid, A. Adaptive LASSO in High-dimensions. https://doi.org/10.31235/osf.io/yphxv (2022).

[5] Ismah, K., Anwar, N. and Bagus, S. (2021). A Multicollinearity-Adjusted Adaptive LASSO for Zero-Infated Count Regression with Weight of Expectation Maximiza- tion Standard Error Adaptive LASSO for Zero Inflated Poisson Data. *Journal Of Physics*, Conference Series. 1776 012050

[6] Tibshirani, R. Regression Shrinkage and Selection via the LASSO. *Journal of the Royal Statistical Society*. **Series B(Methodological)** : 267-288 (1996).

[7] Zou. H, and Hastie. T. "Regularization and Variable Selection via the Elastic net, *Journal of the Royal Statistical Society*.B **67** :301-320 (2005).

[8] Qian, W. and Yang, Y. Model Selection via Standard Error Adjusted Adaptive LASSO. *Ann Inst Stat Math*, **65**:295-318 (2013).

[9] Zou, H. The Adaptive LASSO and its Oracle Properties. *Journal of the American Statistical Association,* **101**: 1418-1429 (2006).

[10] Fan, J. and Lv, J. Sure Independence Screening for Ultrahigh Dimensional Feature Space. *Journal of the Royal Statistical Society B*, **70(5)**: 849-911 (2008).

[11] Buhlmann, P. and Van De Geer, S. *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Springer Science & Business Media, Heidelberg, 2011.

[12] Zou, H. and Zhang, T. On the Adaptive Elastic net with a Diverging Number of Parameters. *Annals of Statistics*. **37**:1733-1751 (2009).

[13] Hastie, T. Tibshirani, R. and Friedman, J. *The Elements of Statistical Learning; Data Mining, Inference and Prediction*. New York, Springer, 2001.

[14] Hastie, T. Tibshirani, R. and Friedman, J. *The Elements of Statistical Learning; Data Mining, Inference and Prediction*. New York, Springer, 2017.

[15] Farhadi, Z., Belaghi, R.A. and Alma, O.G. Analysis of Penalized Regression Methods in a Simple Linear Model on the High-Dimensional Data. *American Journal of Theoretical and Applied Statistics*, **8**(5): 185-192 (2019).

[16] Wu, Y. Cant Ridge Regression Perform Variable Selection? *Technometrics*, **63**(2):263-271 (2021).

[17] Araveeporn, A. The Higher-order of Adaptive LASSO and Elastic net Methods for Classification

[18] Greenwood, C.J. et al. A Comparison of penalized Regression Methods for Informing the Selection of Predictive Markers. *PLoS ONE* 15(11): e0242730. https://doi.org/10.1371/journal.pone.0242 (2020).

[19] Tibshirani, R. , Hastie, T. and Wainwright, M. (2019). *Statistical Learning with Sparsity-The LASSO and Generalizations* .Chapman and hall book, 2019

[20] Golub T. R. et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring, Science. *The Annals of Statistics*, **286**: 531-537 (1999).

[21] Alon, U. et al. (1999). Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proceedings of the National Academy of Sciences. USA* **96**(12): 6745-6750