Improved Lasso (ILASSO) for Gene Selection and Classification in High Dimensional DNA Microarray Data

Isah Aliyu Kargi[1,2]([✉]), Norazlina Ismail[1], Ismail Bin Mohamad[1]
[1] Universiti Teknologi Malaysia, Johor, Malaysia
[2] Nuhu Bamalli Polytechnic, Zaria, Nigeria
isaaliyukargi@gmail.com

Abstract—Classification and selection of gene in high dimensional microarray data has become a challenging problem in molecular biology and genetics. Penalized Adaptive likelihood method has been employed recently for classification of cancer to address both gene selection consistency and estimation of gene coefficients in high dimensional data simultaneously. Many studies from the literature have proposed the use of ordinary least squares (OLS), maximum likelihood estimation (MLE) and Elastic net as the initial weight in the Adaptive elastic net, but in high dimensional microarray data the MLE and OLS are not suitable. Likewise, considering the Elastic net as the initial weight in Adaptive elastic yields a poor performance, because the ridge penalty in the Elastic net grouped coefficient of highly correlated genes closer to each other. As a result, the estimator fails to differentiate coefficients of highly correlated genes that have different sign being grouped together. To tackle this issue, the present study proposed Improved LASSO (ILASSO) estimator which add the ridge penalty to the original LASSO with an Adaptive weight to both $\ell_1 - norm$ and $\ell_2 - norm$ simultaneously. Results from the real data indicated that ILASSO has a better performance compared to other methods in terms of the number of gene selected, classification precision, Sensitivity and Specificity.

Keywords—high dimension data, Penalized Adaptive Elastic net, ILasso, logistic regression, gene selection, cancer classification

1 Introduction

Genetics and molecular biology has been transformed recently as a field of research for genes selection and classification in high dimensional DNA microarray data. In DNA microarray data the number of genes is more than thousands from a hundred or less sample tissue [1], [2]. Because of high dimensionality of DNA microarray cancer data, genes selection and classification become a challenging issue in biomedical research[3], [4]. Classification of cancer and selection of genes have been extensively studied in recent years. Thus, application of penalized logistic regression method aid for the identification of significant genes that are biologically relevant for a specific cancer[5]–[7].

Several methods of genes selection have been suggested to choose small and appropriate genes in high dimensional cancer classification. Some penalized likelihood methods that are capable of performing estimation of model and selection of genes efficiently have been recently considered in [8]–[10]. The most frequent and popular penalized likelihood method is the least absolute shrinkage and selection operator (LASSO) proposed in [11]. The $\ell_1 - norm$ penalty is imposed to the loss function in LASSO. Variable selection can easily be performed by pushing some gene coefficient to exactly zero because of the $\ell_1 - norm$ property[12]. Due to this motive LASSO gained recognition in classification of cancer in high dimensional data.

However, despite the advantageous of LASSO, it has some drawbacks. Firstly, in a scenario of $p > n$, LASSO cannot select variables more than the size of sample (n) as a result of convex optimization problem[13], [14]. Secondly, LASSO is inconsistent in variable selection as it applies the same amount of penalization to all coefficients which resulted to a biased estimation of large coefficients. Thus oracle property is not obtainable[15]. Lastly in cancer classification with high dimensional microarray data, LASSO cannot address the effect of grouping [16], [17]. Furthermore, penalized Elastic net was proposed by Zou and Hastie in [18] which is a mixture of LASSO penalty and Ridge penalty to address some shortcomings of LASSO.

Conversely, like LASSO, the Elastic also has a limitation of not enjoying the oracle property even though it has a better performance than LASSO. To tackle the effect of grouping of highly correlated genes as well as enjoying oracle property in high dimensional data Zou and Zhang proposed Adaptive elastic net in [19]. Though, in the Adaptive elastic net the estimator has influence to the initial weight. Many studies from the literature have proposed the use of ordinary least squares (OLS), maximum likelihood estimation (MLE) and Elastic net as the initial weight in the Adaptive elastic net, but in high dimensional microarray data the MLE and OLS are not suitable [20]. Likewise, considering the Elastic net as the initial weight in Adaptive elastic yields a poor performance, because the ridge penalty in the Elastic net grouped coefficient of highly correlated genes closer to each other. As a result, the estimator fails to differentiate coefficients of highly correlated genes that have different sign being grouped together[21]. To tackle this issue, the present study proposed Improved LASSO (ILASSO) estimator which add the ridge penalty to the original LASSO. Also, in ILASSO, an Adaptive weight is added to both $\ell_1 - norm$ and $\ell_2 - norm$ simultaneously.

The reminder of this paper is organised as follows: introduction is presented in section 1, section 2 displays the penalized Adaptive regression methods and the proposed technique used in the present study. While discussion of results is presented in section 3, and finally section 4 covers the conclusion of the study.

2　Method

2.1　Penalized logistic regression

A statistical method that model a binary classification problem is called logistic regression. The response variable in classification of cancer for logistic regression

method consist of only two values 1 or 0 for tummor and normal class respectively. For binary classification, the regression function has a nonlinear relation with the linear group of genes[22]. Let's denote the response variable for observation i by $y_i \in (0,1), where$ $i = 1,2,3,……,n$. Assume p genes with $i - th$ gene vector represented by $x_i = (x_{1.i}, x_{2,i}, x_{3.i},……,x_{p,i})^T$. Then the linear predictor is given by:

$$L(x_i, \beta) = \beta_0 + \beta_1 x_{1,i} + \cdots + \beta_p x_{p,i} \tag{1}$$

Given the linear predictor, the logit transformation is carried out as follows:

$$Log_e \left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{1,i} + \cdots + \beta_p x_{p,i} \tag{2}$$

The probability of observation i belonging to the positive class, i.e., malignant, can now be calculated as follows:

$$P_i = [y_i = 1|x_i] = \frac{e^{\beta_0+\beta_1 x_{1,i}+\cdots+\beta_p x_{p,i}\cdot}}{1+e^{\beta_0+\beta_1 x_{1,i}+\cdots+\beta_p x_{p,i}\cdot}} = \frac{e^{x_i^T \beta}}{1+e^{x_i^T \beta}} \tag{3}$$

Where, $\beta = (\beta_0, +\beta_1, …\beta_p)^T$ represent the genes coefficient of the unknown vector. Thus, the log-likelihood function of the penalized logistic regression is defined below:

$$LL(\beta) = \sum_{i=1}^n [y_i log(p_i) + (1 - y_i)log(1 - p_i)] \tag{4}$$

Normally cancer data is considered high dimensional data with $p \gg n$. A non- negative penalization term is added to the negative log-likelihood function to control the number of genes coefficients in Eqn. (4). Numerous popular penalization term has been discuss in [11], [18], [23]. The most common penalization term is the $\ell_1 - norm$ proposed by Tibshirani in [11]. In penalized logistic regression (PLR) the $\ell_1 - norm$ is considered as the penalty term that perform gene selection and classification by constraining the negative log-likelihood function for estimation and classification of gene coefficients, that is, $-LL(\beta) + \lambda P(\beta), where$ $\lambda P(\beta)$ is used to penalized the estimated coefficients. The positive turning parameter λ is employed to grip the degree of the shrinkage. When $\lambda = 0$, the penalization term is deserted and PLR abridged to logistic regression with maximum likelihood estimation (MLE). Though, the higher the value of λ the higher the effect on the estimated coefficients. One important part of fitting a model, is the choice of the turning parameter. If the interest is on classification, the turning parameter is usually chosen to minimize the misclassification error by computing the optimal balance where an increase in the variance resulted to a decrease in the bias. Therefore, the objectives function for PLR is describe as:

$$\hat{\beta}_{PLR} = arg\min_{\beta}[\sum_{i=1}^n [y_i log(p_i) + (1 - y_i)log(1 - p_i)] + \lambda P(\beta)] \tag{5}$$

It is assumed without loss of generality that the vectors of genes are standardized, that is.,

$\sum_{j=1}^{n} x_{ij} = 0 \; and \; \frac{1}{n}\sum_{i=1}^{n} x_{ij}^2 = 1, \forall j \in \{1,2,3,.....P\}$. one of the most common PLR methods in the state of art (literature) is LASSO, where the objective function for LASSO is given as follows:

$$\hat{\beta}_{LASSO} = arg \min_{\beta} \left[\sum_{i=1}^{n} [y_i log(p_i) + (1 - y_i) log(1 - p_i)] + \lambda \sum_{j=1}^{p} |\beta_j| \; \right] \quad (6)$$

Note that in LASSO the value of λ i.e, the turning parameter has a high influence on the estimated gene coefficients. As a result, a smaller value of λ reduces some coefficient to zero absolutely. This indicated that LASSO can effectively choose some variables by reducing the coefficient of others to exactly zero. Thus, for features selection in different scientific area LASSO can be used effectively [24], [25]. Method of cross validation is normally employed for selecting the value of λ for identifying the best value of λ that has the smallest measure of accuracy (mis classification error) [26]. Although LASSO has the benefit of performing features selection in higher dimensional data, it has some limitations. In a group of highly correlated variables, LASSO can only choose one variable among the group. Also, LASSO cannot select explanatory variables more than the size of sample (n) and at the same time LASSO does not enjoy the oracle properties.

To tackle these aforementioned shortcomings of LASSO, another penalized likelihood method was proposed in [18], [27]. Elastic net is another penalized regression method which is a combination of $\ell_2 - norm$ and $\ell_1 - norm$ penalties. Elastic net handle the problem that occurs when $\ell_1 - norm$ is used in the scenario of highly correlated variables. Penalized logistic regression using Elastic net is modelled as follows:

$$\hat{\beta}_{Elastic} = arg \min_{\beta} \left[\sum_{i=1}^{n} [y_i log(p_i) + (1 - y_i) log(1 - p_i)] + \lambda \left((1 - \alpha) \sum_{j=1}^{p} |\beta_j| + \alpha \sum_{j=1}^{p} \beta_j^2 \right) \right] \quad (7)$$

It can be seen from Eqn. (7) that two turning parameters are used for Elastic net estimator, i.e., α and λ. Usually, method of cross validation is employed to obtain the optimal value of λ for a fixed value of α

2.2 Penalized Adaptive Logistic Regression

Zou in [23] proposed penalized Adaptive LASSO (ALASSO), to improve the selection consistency of LASSO by introducing a data driven weight to the $\ell_1 - norm$. The idea in the ALASSO is that, by considering moderately large quantity of the shrinkage for the zero coefficient and for the nonzero coefficients small quantity, this may likely decrease estimation biased and enhanced selection of variables stability. Fundamentally, the ALASSO is a convex optimization challenged with ℓ_1 restriction. Therefore, the efficient algorithm for LASSO solution can be employ as well to solve ALASSO. The penalized Adaptive LASSO is therefore defined as:

$$\hat{\beta}_{ALASSO} = arg \min_{\beta} \left[\sum_{i=1}^{n} [y_i log(p_i) + (1 - y_i) log(1 - p_i)] + \lambda \sum_{j=1}^{p} w_j |\beta_j| \right] \quad (8)$$

Where w_j represent the weight for the $j - th$ coefficient and is computed as; $w_j = \left(|\hat{\beta}_j| \right)^{-\gamma}$, such that $\hat{\beta}$ is the initial value of the estimator and $\gamma > 0$. The true value of

the estimate converge with $O\left(n^{-1/2}\right)$ [20]. Adaptive LASSO has a drawback of not handling the group of highly correlated genes in high dimensional cancer data. As a result, Elastic net was proposed to address the effect of grouping in highly correlated variables as stated in [28]. Despite the effectiveness of Elastic net in dealing with grouping effect in highly correlated variables, the Elastic net does not possess oracle property[20]. Hence, the Adaptive Elastic net was proposed in [19], [29] which is presented mathematically as:

$$\hat{\beta}_{AElastic} = arg\min_{\beta}\left[\sum_{i=1}^{n}[y_i log(p_i) + (1 - y_i)log(1 - p_i)] + \lambda\left((1 - \alpha)\sum_{j=1}^{p} w_j |\beta_j| + \alpha\sum_{j=1}^{p}\beta_j^2\right)\right] \qquad (9)$$

2.3   Proposed method

Penalized Adaptive Elastic net is consistent in selection of variables as well as handling the effect of grouping and has the oracle property. But the initial weight has influence on the performance of the estimator. For estimating the initial weight to the Adaptive Elastic net, existing works have proposed the used of maximum likelihood estimation (MLE) and ordinary leas squares (OLS). However, for data of lower dimensional, MLE and OLS may be appropriate, but for classification of cancer in high dimensional data they have a poor performance [19], [29], [30]. Equally, considering Elastic net as the initial weight is not suitable due to the fact that, the ridged penalty in the Elastic net grouped the coefficients of correlated genes closer to each other. Though, as a result the estimator fails to distinguish significantly highly correlated genes that have different signs [19]. This problem resulted in genes with different features to be grouped together. In the present paper, ILASSO was proposed to address this problem by introducing the ridge penalty to the novel LASSO estimator. As a result, the estimator can easily identify group of highly correlated genes with different magnitudes.

The objective function for ILASSO is given as follows:

$$\hat{\beta}_{ILASSO} = arg\min_{\beta}\left[\sum_{i=1}^{n}[y_i log(p_i) + (1 - y_i)log(1 - p_i)] + \lambda\left(\sum_{j=1}^{p} W_j\left((1 - \alpha)|\beta_j| + \alpha\beta_j^2\right)\right)\right] \qquad (10)$$

Note that in ILASSO, the adaptive weight is applied to the LASSO penalty and ridge penalty respectively. The turning parameter α is titled towards the ridge penalty to achieved the effect of grouping. Furthermore, the Adaptive weight in ILASSO are chosen in an exclusive manner to take care of coefficient of highly correlated genes, in other word, the initial weight is generated using LASSO estimator. For the estimator to maintain the grouping effect of highly correlated genes the value of α is considered to be $\alpha > 0.5$. However, applying LASSO estimator as the initial weight in the ILLASSO, we attain a shrinkage so that coefficient of highly correlated genes especially those with different magnitude will be pushed to exactly zero because of the higher value of α .

3   Discussion of results

The proposed ILASSO method has been evaluated using a well-known microarray data, kwon as Diffuse Large B-cell lymphoma (DLBCL) available in [31]. The dataset

consists of 77 samples of genes expression values that are measured with high density oligonucleotide microarrays of the most two prevalent adult lymphoid malignancies which comprise 58 samples of diffuse large b-cell lymphomas and 19 samples of follicular lymphoma (FL).

For purpose of comparison between the propose ILASSO and other methods DLBCL dataset is divided into testing and training data set. The training data set was generated randomly with 70% of the original data whereas 30% of the remaining data set was used as the testing data. Method of 10-fold cross validation (CV) was employed to obtain the optimal value of λ. The study apply a turning grid in obtaining the optimal of α where the value of α is incremented in steps and in each iterations the resulting performance metrics is presented. Then for evaluating an estimator the value of α that gives the best result is considered. All computations in this paper was carry out using R software.

Similarly, number of gene selected (#G), the Area under the curve (AUC), Sensitivity (Sen), Specificity (Spe) and informediness (IF) are the performance metric used for the study. One major important factor in higher dimensional DNA microarray data is the number selected by each techniques [32] [33], [34]. Simple and interpretable model is not available by the method that selected higher number of genes.

From Table 1, it can be observed clearly that the proposed ILASSO method outperform the other four methods considering the number of genes chosen, ILASSO selected 6 number of genes compared to LASSO, ALASSO, Elastic net and Adaptive Elastic net that obtain 29, 12,50 and 14 respectively. Also for classification accuracy, from Table 1, ILASSO is superior with the highest value of AUC 0.99 which is much better than 0.930, 0.980, and 0.960 attained by LASSO, ALASSO, Adaptive Elastic net and Elastic net. In addition, ALASSO and Adaptive approaches with a specificity from table 1, ILASSO is considered superior to the other methods. For example,the proposed ILASSO has the highest sensitivity of 95%. This indicated that ILASSO significantly outperformed the other methods in classifying patients having DLBCL cancer with a probability of 0.95. Likewise, outcomes for specificity from table 1 which stand for the probability of an adaptive penalized techniques in classifying normal patients. It can be seen in terms of specificity that ILASSO and ALASSO has a better performance with a sure probability of 1compared to LASSO, Elastic net and Adaptive Elastic net. This might be as a result of be as a result of the new weight added to both LASSO and Ridge penalty.

To further evaluated the performance of the proposed ILASSO, an alternative measure of imbalance data call informedness (IF) was used. We can observe from Table 1 that ILASSO outperforms the other with an IF of 0.952. However, this shows the ability of our proposed technique in measuring and evaluating predictive models for DLBCL cancer data set.

Conversely, Fig. 1 –3 displays the 95% confidence interval for the mean of AUC, Sensitivity and Specificity for DLBCL cancer data. It can be clearly seen from Figure 1 – 3 that ILASSO emerged as the best classifier with smaller variance for all the performance metrics. Moreover,ILASSO and ALASSO has similar results of classifying patients having DLBCL cancer with a xure probability of 1 over the 50 partition metrics.

Table 1. Classification performance results on DLBCL cancer data.

Method	#G	AUC	Sen.	Spec.	IF
LASSO	29	0.930	0.929	0.833	0.762
ALASSO	12	0.980	0.905	1.00	0.905
Elastic Net	50	0.960	0.929	0.917	0.845
Adaptive Elastic Net	14	0.980	0.929	0.980	0.929
FAElastic	6	0.990	0.952	1.00	0.952
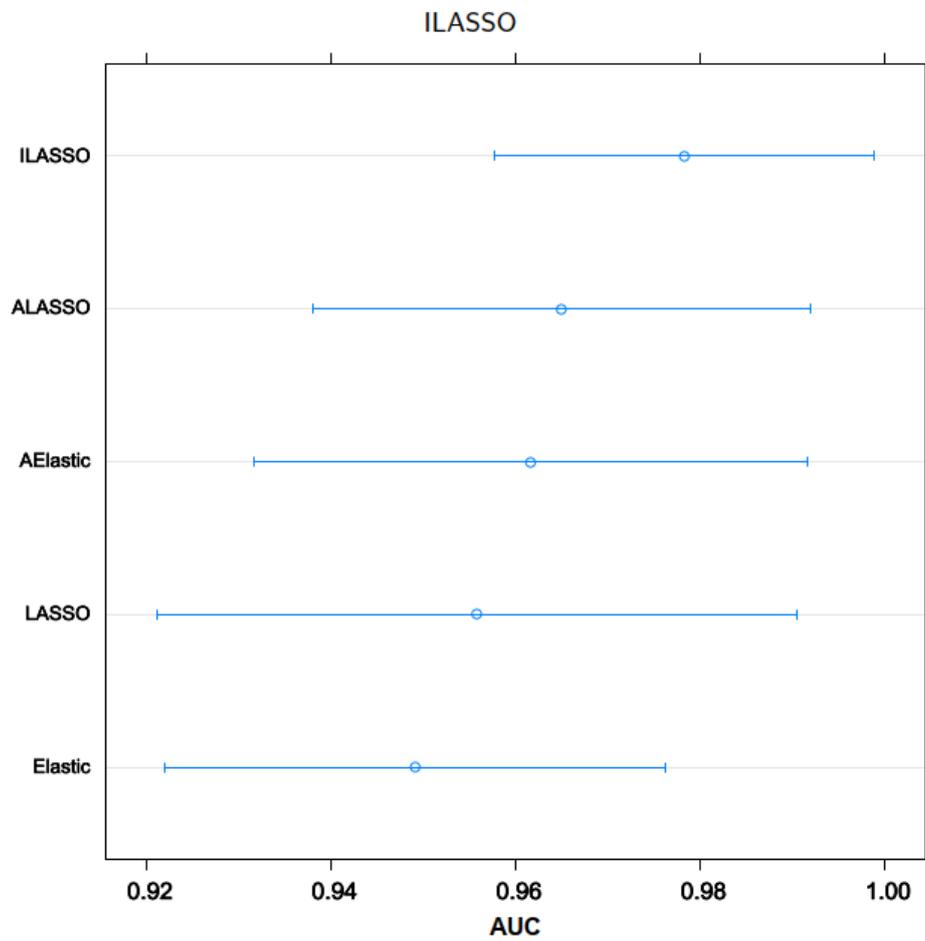
Fig. 1. Performance comparison of ILASSO with existing methods on the basis of average AUC:95% confidence interval.
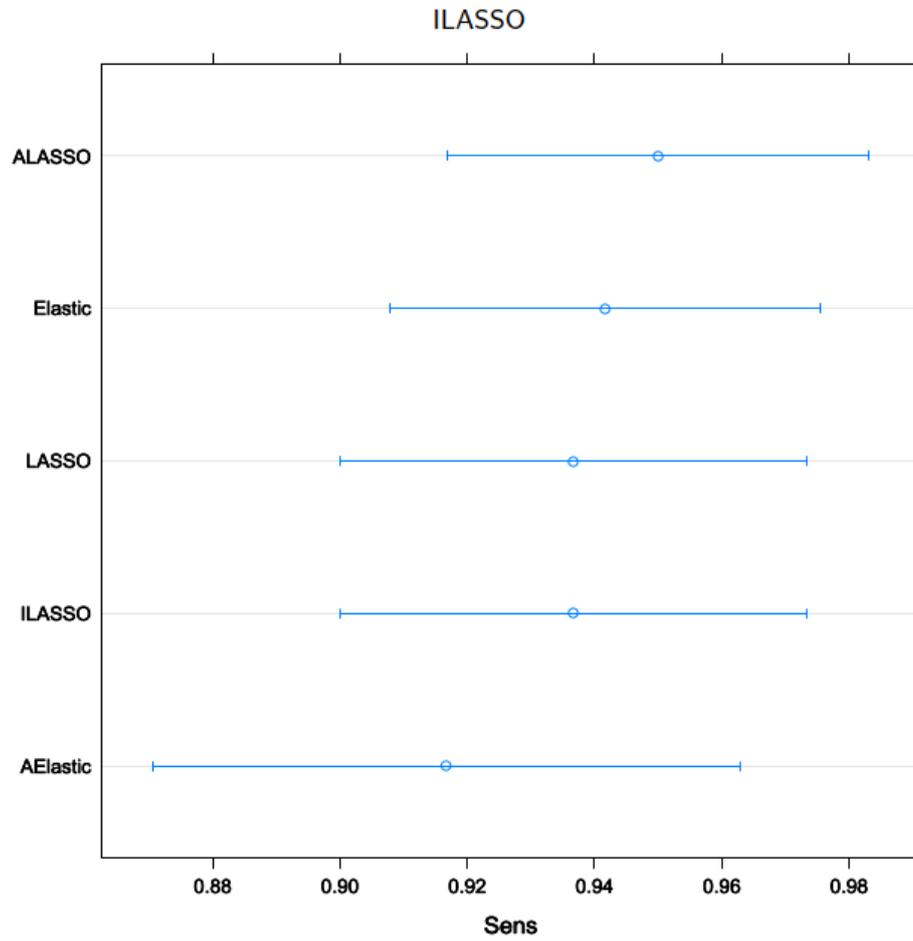
Fig. 2. Performance comparison of ILASSO with existing method on the basis of average sensitivity: 95% confidence interval.
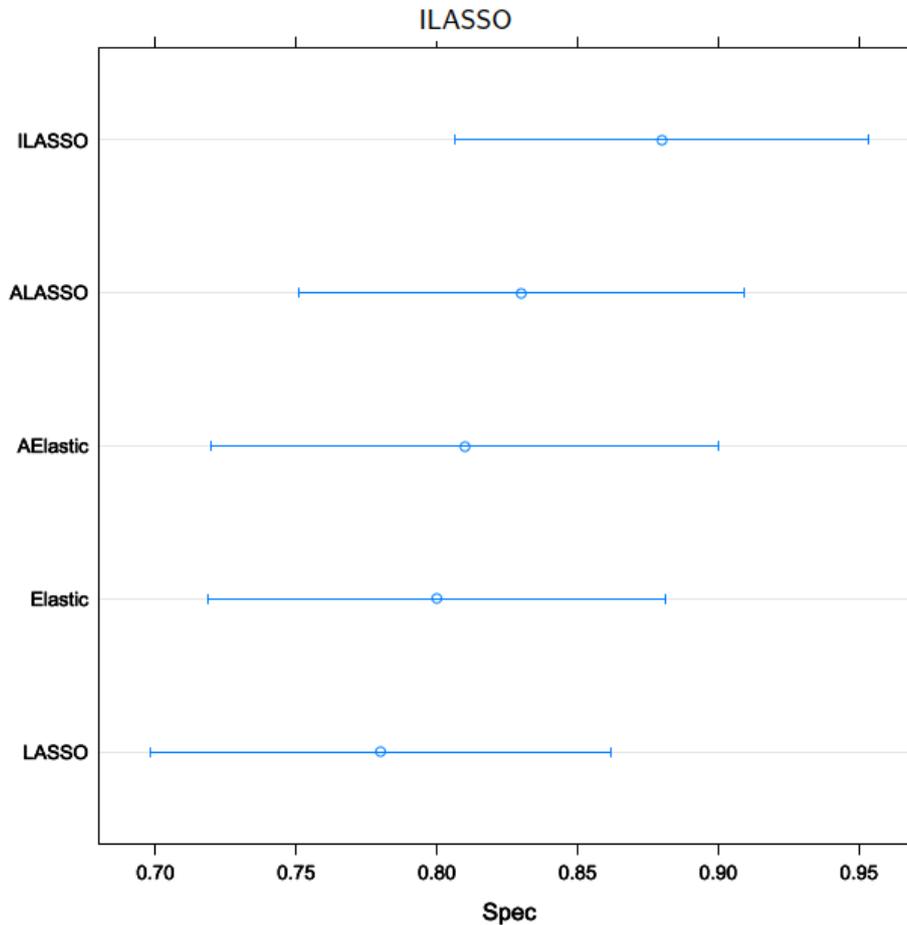
Fig. 3. Performance comparison of ILASSO with existing methods on the basis of average specificity:95% confidence intervals.

4   Conclusion

Classification of cancer is one of the most significant area of research in high dimensional DNA microarray data. However, various computational methods are unsuccessful in selecting and classifying small subsets of significant genes due to problem of high dimensionality of cancer data. To tackle this problem, the current study proposed ILASSO penalized logistic regression method for gene selection and classification in high dimensional data. Results from the study has confirmed the superiority of the proposed ILASSO method in terms of number of genes selected, classification precision, Sensitivity and Specificity.

Overall, the results indicated that ILASSO is capable of handling and analysing high dimensional data correctly. Furthermore, the proposed ILASSO results can be useful

for classification of cancer and other related high dimensional data. Hence, the proposed ILASSO technique can be practically employed in medical research area.

5   Conflict of Interest

The authors declare that there is no conflict of interest associated with this manuscript.

6   Acknowledgements

7   Reference

[1] S. Zheng and W. Liu, "An experimental comparison of gene selection by Lasso and Dantzig selector for cancer classification," *Comput. Biol. Med.*, vol. 41, no. 11, pp. 1033–1040, 2011. https://doi.org/10.1016/j.compbiomed.2011.08.011

[2] X. Liu, S. Wang, H. Zhang, H. Zhang, Z. Yang, and Y. Liang, "Novel regularization method for biomarker selection and cancer classification," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. PP, no. c, p. 1, 2019. https://doi.org/10.1109/tcbb.2019.2897301

[3] T. Basu and J. Einbeck, "Binary Credal Classification under Sparsity Constraints," no. January, 2020.

[4] M. Ijaz, Z. Asghar, and A. Gul, "Computation Ensemble of penalized logistic models for classification of high-dimensional data," *Commun. Stat. - Simul. Comput.*, vol. 0, no. 0, pp. 1–17, 2019. https://doi.org/10.1080/03610918.2019.1595647

[5] S. Wu, H. Jiang, H. Shen, and Z. Yang, "Gene Selection in Cancer Classification Using Sparse Logistic Regression with L1/2 Regularization," *Appl. Sci.*, vol. 8, no. 9, p. 1569, Sep. 2018. https://doi.org/10.3390/app8091569

[6] S. Kar, K. Das Sharma, and M. Maitra, "Gene selection from microarray gene expression data for classification of cancer subgroups employing PSO and adaptive K-nearest neighborhood technique," *Expert Syst. Appl.*, vol. 42, no. 1, pp. 612–627, 2015. https://doi.org/10.1016/j.eswa.2014.08.014

[7] G. Sun, J. Li, J. Dai, Z. Song, and F. Lang, "Feature selection for IoT based on maximal information coefficient," *Futur. Gener. Comput. Syst.*, vol. 89, pp. 606–616, 2018. https://doi.org/10.1016/j.future.2018.05.060

[8] Y. W. Lin, N. Xiao, L. L. Wang, C. Q. Li, and Q. S. Xu, "Ordered homogeneity pursuit lasso for group variable selection with applications to spectroscopic data," *Chemom. Intell. Lab. Syst.*, vol. 168, no. January, pp. 62–71, 2017. https://doi.org/10.1016/j.chemolab.2017.07.004

[9] C. Cui and D. Wang, "High dimensional data regression using Lasso model and neural networks with random weights," *Inf. Sci. (Ny).*, vol. 372, pp. 505–517, 2016. https://doi.org/10.1016/j.ins.2016.08.060

[10] M. Dashtban and M. Balafar, "Genomics Gene selection for microarray cancer classi fi cation using a new evolutionary method employing arti fi cial intelligence concepts," *Genomics*, vol. 109, no. 2, pp. 91–107, 2017. https://doi.org/10.1016/j.ygeno.2017.01.004

[11] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso Robert Tibshirani," *J. R. Stat. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

[12] Z. Y. and H. T. M. A. Algamal, "An efficient gene selection method for high-dimensional microarray data based on sparse logistic regression," no. April, 2017.

[13] K. You, S. Member, R. Tempo, and P. Xie, "Distributed Algorithms for Robust Convex Optimization via the Scenario Approach," vol. 9286, no. c, pp. 1–16, 2018.

[14] B. Peter, "High-dimensional statistics , with applications to genome-wide association studies," 2010.

[15] S. V. D. G. Karim Lounici, Massimiliano Pontil, and A. B. Tsyvakov, and Tsybakov, "Oracle inequalities and optimal inference under group sparsity 1," vol. 39, no. 4, pp. 2164–2204, 2011. https://doi.org/10.1214/11-aos896

[16] R. Alhamzawi and H. T. M. Ali, "The Bayesian adaptive lasso regression," *Math. Biosci.*, vol. 303, no. January, pp. 75–82, 2018. https://doi.org/10.1016/j.mbs.2018.06.004

[17] M. Dashtban, M. Balafar, and P. Suravajhala, "Genomics Gene selection for tumor classi fi cation using a novel bio-inspired multi- objective approach," *Genomics*, vol. 110, no. 1, pp. 10–17, 2018. https://doi.org/10.1016/j.ygeno.2017.07.010

[18] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," pp. 301–320, 2005.

[19] S. Ghosh, "On the grouped selection and model complexity of the adaptive," pp. 451–462, 2011.

[20] Z. Y. Algamal and M. H. Lee, "Regularized logistic regression with adjusted adaptive elastic net for gene selection in high dimensional cancer classification," *Comput. Biol. Med.*, vol. 67, pp. 136–145, 2015. https://doi.org/10.1016/j.compbiomed.2015.10.008

[21] J. Wang. S., Nan, B., Rosset, S., Zhu, "Random lasso," vol. 5, no. 1, pp. 468–485, 2011.

[22] H. H. Huang, X. Y. Liu, and Y. Liang, "Feature selection and cancer classification via sparse logistic regression with the hybrid L1/2 +2 regularization," *PLoS One*, vol. 11, no. 5, 2016. https://doi.org/10.1371/journal.pone.0149675

[23] H. Zou, T. Hastie, and R. Tibshirani, "Sparse principal component analysis," *J. Comput. Graph. Stat.*, vol. 15, no. 2, pp. 265–286, 2006.

[24] Z. Yahya, R. Alhamzawi, H. Taha, and M. Ali, "Gene selection for microarray gene expression classi fi cation using Bayesian Lasso quantile regression," *Comput. Biol. Med.*, vol. 97, no. April, pp. 145–152, 2018. https://doi.org/10.1016/j.compbiomed.2018.04.018

[25] W. Rejchel, "Neurocomputing Oracle inequalities for ranking and U -processes with Lasso penalty," vol. 239, pp. 214–222, 2017. https://doi.org/10.1016/j.neucom.2017.02.018

[26] N. A. Al-Thanoon, O. S. Qasim, and Z. Y. Algamal, "Tuning parameter estimation in SCAD-support vector machine using firefly algorithm with application in gene selection and cancer classification," *Comput. Biol. Med.*, vol. 103, no. October, pp. 262–268, 2018. https://doi.org/10.1016/j.compbiomed.2018.10.034

[27] D. X. Zhou, "On grouping effect of elastic net," *Stat. Probab. Lett.*, vol. 83, no. 9, pp. 2108–2112, 2013.

[28] J. Park, "Tolerance intervals from ridge regression in the presence of multicollinearity and high dimension," *Stat. Probab. Lett.*, vol. 121, pp. 128–135, 2017. https://doi.org/10.1016/j.spl.2016.10.016

[29] H. Z. A. H. H. Zhang, "On the Adaptive Elastic-net with a diverging number of Parameters," vol. 37, no. 4, pp. 1733–1751, 2009.

[30] S. Dasgupta, K. Khare, and M. Ghosh, "Asymptotic expansion of the posterior density in high dimensional generalized linear models," *J. Multivar. Anal.*, vol. 131, pp. 126–148, 2014. https://doi.org/10.1016/j.jmva.2014.06.013

[31] T. et al Shipp, M. A., Ross, K. N., Tamayo, P., Weng, A. P., Kutok, J. L., Aguiar, R. C., "Diffuse large B-cell lymphoma outcome prediction by gene- expression profiling and supervised machine learning," vol. 8, no. 1, pp. 2–8, 2002. https://doi.org/10.1038/nm0102-68

[32] C. Kang, Y. Huo, L. Xin, B. Tian, and B. Yu, "Feature selection and tumor classification for microarray data using relaxed Lasso and generalized multi-class support vector machine," *J. Theor. Biol.*, vol. 463, pp. 77–91, 2019. https://doi.org/10.1016/j.jtbi.2018.12.010

[33] Z. Y. Algamal and M. H. Lee, "Penalized logistic regression with the adaptive LASSO for gene selection in high-dimensional cancer classification," vol. 42, pp. 9326–9332, 2015. https://doi.org/10.1016/j.eswa.2015.08.016

[34] Z. Y. Algamal and M. H. Lee, "A two-stage sparse logistic regression for optimal gene selection in high-dimensional microarray data classification," *Adv. Data Anal. Classif.*, vol. 13, no. 3, pp. 753–771, 2019. https://doi.org/10.1007/s11634-018-0334-1

8   Authors

Isah Aliyu Kargi is with Universiti Teknologi Malaysia, Department of Mathematics, Faculty of Science, Johor, Malaysia; and Nuhu Bamalli Polytechnic, Department of Mathematics and Statistics, Zaria, Nigeria.

Norazlina Ismail is with Universiti Teknologi Malaysia, Department of Mathematics, Faculty of Science, Johor, Malaysia.

Ismail Bin Mohamad is with Universiti Teknologi Malaysia, Department of Mathematics, Faculty of Science, Johor, Malaysia.