International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2024; 9(1): 83-89

https://www.mathsjournal.com
Received: 08-11-2023
Accepted: 13-12-2023

Wakaa Ali Hadba
Teaching, Department of Statistics, College of Administration and Economics, University of Kirkuk, Iraq

Hadeel Imad Naser
Teaching, Department of Statistics, College of Administration and Economics, University of Kirkuk, Iraq

Corresponding Author:
Wakaa Ali Hadba
Teaching, Department of Statistics, College of Administration and Economics, University of Kirkuk, Iraq

Comparison of lasso logistic regression, artificial neural networks, and support vector machine in predicting breast cancer

Wakaa Ali Hadba and Hadeel Imad Naser

Abstract
In recent times, breast cancer has surpassed all other types of cancer to become the most widespread and primary cause of mortality among women globally. This study aims to forecast the benign or malignant nature of a breast tumor by employing various machine learning methods, including Support Vector Machines (SVM), Lasso Logistic Regression (LLR), Artificial Neural Networks (ANN), and Logistic Regression (LR). The findings of this research could potentially assist oncologists in accurately diagnosing the specific type of breast tumor. Various performance metrics were employed to assess the efficacy of training and validated models, including Accuracy or Classification Error, Sensitivity, Specificity, and Receiver Operating Characteristic (ROC). The models were constructed and evaluated to determine the optimal performing model. A dataset separate from the one used for model development was employed to validate each model. The data analysis results indicate that the Lasso Logistic Regression (LLR) model outperforms other models in classifying breast cancers. It exhibited superior performance in terms of accuracy, classification error, sensitivity, specificity, and ROC. Furthermore, it mitigates the issues of multicollinearity and high-dimensionality.

Keywords: Logistic regression (LR), lasso logistic regression (LLR), artificial neural networks (ANN), and support vector machine (SVM)

1. Introduction

Cancer cells are cells that develop and replicate rapidly to create tumors (either benign or malignant) and have the potential to spread to other body organs and create new tumors. The second most common reason why the number of female deaths has increased is that this disease has a more significant impact on women than on men. There have been 2.3 million diagnoses. In 2020, there will be 685,000 breast cancer-related deaths worldwide and 7.8 million women who have received a diagnosis during the previous five years, increasing the prevalence of the disease globally. Breast cancer begins in the feeding tissues of the breast and develops in the cell lining (the epithelium) of the ducts by 85% and the sternal by 15%. Breast cancer can have a wide range of causes. It could be hereditary (meaning that a family member or close relative is affected), age (since infection risk rises with age), postmenopausal obesity or excess weight, alcohol or smoking, or other factors. Early intervention increases the likelihood of receiving therapy, increasing survival [1].

Different methods and techniques were employed by numerous researchers who studied breast cancer to identify and predict breast cancer.

(Ohno-Machado & Bialek) In (1998), they used Logistic Regression (LR) and Neural Networks (NN) to compare the factors they chose to create a classification model for the diagnosis of breast cancer. To construct these models, they examined 460 patients and relied on nine pathological features. LR compared the variables. They concluded that both models have similarities [2].

In 2004, Abdolmaleki *et al.* developed two models to distinguish between benign and malignant breast tumors. One model was based on an artificial neural network (ANN), while the other model used logistic regression analysis (LRM). The researchers compared the performance of these models in differentiating tumors using the medical records of 161 patients.

Researchers compared the two models using ROC analysis and found that the trained network performed better. An average ANN had 98% sensitivity, 90% accuracy, and 67% specificity, while in the LRM, the outcome indicated an increase from 60% to 93%, keeping an accuracy around 90% [3].

(Ayer *et al.*) in (2010) examined and compared the LR and ANN models and their application to estimate the risk of breast cancer based on demographic risk factors and mammographic descriptors. They demonstrated that the two models performed similarly [4].

(Faradmal *et al.*) in (2014) conducted a historical cohort study of 104 breast cancer patients from 1997 to 2005. The area under the curve (AUC) for artificial neural network (ANN) in the first, second, and third years after diagnosis were 0.918, 0.780, and 0.800, respectively. For linear logistic model (LLM), the AUCs were 0.834, 0.733, and 0.616 in the same time periods. The mean AUC for ANN was much higher compared to LLM. (0.845 vs. 0.744), and the researchers concluded that the prediction ability in ANN is higher than in LLM [5].

(Khosravanian & Ayat) In (2016), they developed a Decision Support System (DSS) using a probabilistic neural network (pnn) to identify the type of breast cancer in patients. The data used to analyze the proposed method contained 699 cases of BC patients. Since the Sensitivity was 1, the Specificity was 0.98, and the accuracy was 0.99, they concluded that the system's proposed performance based on the three indicators in the network test was satisfactory [6].

(Rajbharath *et al.*) (2017) proposed combining logistic regression and random forest algorithms to create a hybrid breast cancer survival prediction model. The logistic regression approach was used to collect data from the WDBC dataset. Sensitivity, Specificity, and accuracy were utilized to assess the collecting method, and it was discovered that the suggested method offers greater accuracy [7].

(Nourelahi *et al.*) in (2019), A proposed methodology aims to predict a patient's survival for a period of 60 months following a breast cancer diagnosis. The model was constructed using data from 5673 breast cancer patients obtained from the Breast Disease Research Centre at Shiraz University of Medical Sciences in Shiraz, Iran. The researchers constructed a model utilizing 1930 occurrences and 16 features through the use of logistic regression. Subsequently, they selected the model that exhibited the highest levels of accuracy, specificity, and sensitivity (72.49%, 72.83%, and 71.85%) [8].

(Islam *et al.*) 2020) examined the performance of five supervised machine learning methods: Random Forest, ANN, LR, K-Nearest Neighbours, Support Vector Machine (SVM), and ANN. We obtained the data from the UCI machine learning database. Several metrics were used to quantify performance, including F1 score, accuracy, Sensitivity, Specificity, precision, negative predictive value, and false positive and false negative rates. In contrast to ANN, which had accuracy, prediction, and F1 score of 98.57%, 97.82%, and 0.9890, SVM had these values (97.14%, 95.65%, and 0.9777) [9].

In 2022, (ÖZMEN-AKYOL) created (ANN and LR Logistic Regression Models) to predict breast cancer tumors using 699 examples and ten characteristics for training and testing. The data was organized into nine input parameters and an output parameter. The accuracy of the LR model is better than that of the ANN at 96.1% [10].

(Ebrahim *et al.*) in (2023) developed classification models using Machine Learning (ML) binary classification to distinguish between benign and malignant breast cancer tumors. The data provider, which included 1.7 million records, was the National Cancer Institute (NIH), USA. Utilized were ensemble techniques (ET), classical decision tree (DT) algorithms, linear discriminant (LD) support vector machine (SVM), and logistic regression (LR). They also employed deep and recurrent neural networks and the Probabilistic Neural Network (PNN) to compare. Attaining a 98.7% accuracy rate, the findings demonstrated The decision trees and ensemble strategies outperform the other methods [11].

In order to determine if a breast tumor is benign or malignant, this study used a range of machine learning approaches, such as support vector machines (SVM), ANNs, Lasso Logistic Regression (LLR), and logistic regression (LR), and compare them to find the model that would provide the best performance. This paper is divided into four sections: the first includes an introduction and literature review in the area of breast cancer, the second section includes the used methodology, the third presents data collection and Result and discussion, and finally, section four include a conclusion.

2. Materials and Methods

2.1 Logistic Regression (LR)

It is a particular case of general linear regression. It involves the analysis of descriptive data [12]. LR does not impose any restrictions on the independent variables, and it can be descriptive, continuous, discrete, or a mixture of quantitative and descriptive variables [13]. Binary logistic regression is utilized when the dependent variable has only two possible values that follow a Bernoulli distribution. The likelihood that the response (success) will occur is valued at 1, and the probability that the response (failure) will not occur is valued at 0 [14]. LR gets its name from the logit transform applied to Y. Mathematically, LR depends on probabilities, odds and log odds since odds = $\frac{\pi}{1-\pi}$ (1), by taking a natural log, then the logit function can be written with the following expression :

$$\ln\left(\frac{\pi}{1-\pi}\right) = \hat{b}_0 + \hat{b}_1 x_1 + \cdots + \hat{b}_j x_n \tag{2}$$

Equation (2) can be written in the following form, which is called the logistic regression model:

$$p(x) = \frac{1}{1+e^{-\sum \hat{b}_j x_{ij}}} \tag{3}$$

Where i=1,…., n, j=1,…,k, p(x) the value of the output (dependent variable), e the natural logarithm, x_{ij} the independent variables and $\hat{b}_0, \dots, \hat{b}_j$ the model coefficients [15, 16].

2.2 The Lasso Logistic Regression (LLR)

A general form methodology for finding coefficient estimators for the linear method shown below is the Lasso (Least Absolute Shrinkage and Selection Operator) regression method for shrinking the predictors' coefficients and penalizing the coefficient by shrinking towards zero or close to zero: [17]

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{N} \left(y_i - (X\beta)_i\right)^2 \tag{4}$$

Moreover, the lasso estimator is given by

$$\min_{\beta \in \mathbb{R}^p} \left\{ \frac{1}{2N} \parallel y - X\beta \parallel_2^2 + \lambda \parallel \beta \parallel_1 \right\} \quad (5)$$

Where: $\parallel \beta \parallel_1 = \sum_{i=1}^{N} |\beta_i|$ In the lasso method obtained by generalized cross-validation, the tuning parameter λ>0 regulates the penalty's strength, referred to as the shrinkage penalty. A characteristic of Lasso is the ability to choose variables from a vast set. As a result, Lasso performs regularization and variable selection. Due to the non-smooth nature of the constraint, the lasso problem's solutions are nonlinear in y_i [18, 19].

2.3 Artificial Neural Networks (ANN)

It is a simulation of biological neural networks that analyzes data similarly to the human brain [20]. The ANN consists of a large group of processing elements called neurons (nodes), connected by links or communication links, that pass the input signals to other neurons in other layers [21]. Since each neuron receives information from several neurons through connections, the collection process takes place by multiplying the value of the input Xi by its weights. θ_i (each neuron has a certain weight) and then collecting it in the form: $u = \sum_{i=1}^{n} \theta_i x_i + \theta_0$ (6) where i=1,…,n and that $\theta_0$is the bias value, which is a constant value equal to 1. Each neuron in an ANN is processed by a nonlinear function known as the activation or transfer function during the assembly step, and the result of this process is represented by the letter a, with the formula a = f (u) (7) Where a represents the value of a new entry for the neuron in the other layer. F is the activation function [22, 23]. Up to the neuron of the ANN's last layer is the output Y (X), which may be binary and takes values 0 and 1. Single-layer neural networks have input and output layers, and multicellular neural networks, which have input and output layers and are connected by one or more hidden layers, are the two types of neural networks [24]. The most essential characteristic of ANN is its ability to simulate the behavior of the biological nervous system, such as learning and training, which enables it to remember knowledge [25]. Then, the ANN is trained by operating an error backpropagation (algorithm) in which the cells are connected towards the front, and then the process returns towards the back. This process continues; modify the cell weight values until the ANN output has the most negligible error possible [26].

2.4 Support Vector Machine (SVM)

SVM is a high-dimensional learning method that uses a fictitious linear function. Statistical theory-derived learning bias is used to train SVM through techniques based on optimization theory. The principal objective of this methodology is to generate an OSH (Optimal Separating Hyperplane), which yields the optimal separation function for classification [9]. The fields that effectively divide data are those with the highest margins and adequate data separation. Two classes can be separated by a pair of parallel bounding planes, data located in the boundary field called the support vector:

First-class constraints are placed by the first delimiter field, and second-class constraints are placed by the second boundary field.

$$\begin{aligned} x_i w + b &\geq +1, y_i = +1 \\ x_i w + b &\leq -1, y_i = -1 \end{aligned} \quad (8)$$

The position of the alternate field relative to the coordinate center, or the distance between the bounding plane, is represented by the variable b and is calculated using the spacing-to-center formula. W is the conventional field. The optimal value of this margin is achieved when equation (8) is satisfied. The values of B and W will be multiplied by a constant to provide a margin value that is also multiplied by the same constellation. Therefore, by adjusting the values of b and w, the constraint in the equation can be satisfied. In order for the inequality to accurately depict the two limiting boundaries of equation (4), minimize ‖w‖^2 equal to maximize 1/‖w‖.

$$y_i(x_i w + b) - 1 \geq 0 \quad (9)$$

An optimization problem with constraints can be developed for the field that has the most optimal division and yields the biggest profit margin, that is, min $\frac{1}{2}||w||^2$ (10) Using a Lagrange multiplier, This problem can be solved earlier if transformed into a Lagrange formula. Consequently, the constraint optimization problem can become:

$$\min_{w,b} L_p(w, b, \alpha) = \frac{1}{2} \parallel w \parallel^2 - \sum_{i=1}^{n} \alpha_i y_i (x_i w + b) + \sum_{i=1}^{n} \alpha_i \quad (11)$$

By adding constraint, Minimizing Lp to w and b, the equation (11) is obtained :

$$w = \sum_{i=\pm}^{n} \alpha_i y_i x_i \text{ and } \sum_{i=1}^{n} \alpha_i y_i = 0 \quad (12)$$

Lagrangian formula LP is transformed into LD (dual problem) for the primary problem. In order to find the specific coefficients of the optimal hyperplane defined by (w0, b0), we manipulate the equations to get: [23]

$$L_D = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j x_i x_j \quad (13)$$

2.5 Performance evaluation of models

There are various metrics for the evaluation of models

Table 1: Confusion matrix

Observed	Predicted		Total
	Positive	Negative	
Positive	TP	FN	P
Negative	FP	TN	P'
Total	q	q'	1

Sensitivity (SE) (the value of the probability that the expected classification will be correct for the case that it is correct) is calculated: $SE = \frac{TP}{TP+FN} = \frac{TP}{P}$ (True positive) TP: The Number of cases that are classified as accurate and are, in fact, true.

(False negative) FN: The Number of cases that were classified as false and are, in fact, true.

Specificity (SP) (the value of the probability that the expected classification will be wrong for the wrong case) is calculated:

$$SP = \frac{TN}{FP + TN} = \frac{TN}{P'}$$

TN: The Number of cases that are classified as false and are, in fact, false.
FP: The Number of cases that are classified as accurate and are, in fact, false [27].
Accuracy (AC) (the ratio of correctly classified to all cases and calculated: $AC = \frac{TP+TN}{TP+TN+FP+FN}$ Another important criterion used to determine the efficiency of the classifier is the ROC curve (the area under the receiver operating characteristic) through which 1- Specificity is drawn on the x-axis and Sensitivity on the y-axis, giving the area under the curve that ranges from 0 to 1 is a measure of the model's ability to distinguish between cases [28].

3. Results and Discussion

Breast cancer is the most common type of cancer in women worldwide. It accounted for 25% of all cancer diagnoses and impacted approximately 2.1 million people in 2015 alone. It starts when the growth of breast cells becomes uncontrollable. These cells usually grow into tumors that appear on X-rays or can be felt as lumps in the breast area. Determining a tumor's malignant (cancerous) or benign (non-cancerous) nature is the primary barrier to its identification. Features are calculated from a digital picture of a fine needle aspirate (FNA) taken of a breast mass. In the three-dimensional space outlined in [K. P. Bennett and O. L. Mangasarian: "Robust Linear Programming Discrimination of Two Linearly Inseparable Sets," Optimisation Methods and Software 1, 1992, 23–34], discuss the properties of the cell nuclei in the image.
This database is also available through the UW CS ftp server:
Ftp ftp.cs.wisc.educated math-prog/cpo-dataset/machine-learn/WDBC
Also can be found on Kaggle:
https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data
It also can be found on the UCI Machine Learning Repository:
https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29.

Attribute Information

ID number, the predictive variables are (30), the dependent variable is the diagnosis (M = malignant, B = benign), all feature values are recorded with four significant digits, and there are no missing values (missing attribute values: none). The information included 357 benign and 212 malignant class distributions, 569 instances, and 32 attributes (ID, diagnosis, and 30 real-valued input features). We denote it with symbol and symbol (B, M). The ratio of B in the data is (0.63), and the ratio of M is (0.37). For analysis, the data were divided into a ratio (0.8) for training and a ratio (0.2) for testing, The Number of views for training (456) and the Number of views for testing (113).

Table 2: ROC of the models

ROC	Min.	1st Qu.	Median	Mean	3rd Qu.	Max
logisic	0.90170	0.925955	0.948916	0.943974	0.967546	0.97574
lasso	0.9639	0.99742	0.99742	0.991561	0.999484	0.9994
svm	0.9852	0.989164	0.991228	0.991280	0.992776	0.9979
Neural_with_LDA	0.9703	0.990196	0.994324	0.990868	0.999484	1

Table 3: Sensitivity of models

Sensitivity	Min.	1st Qu.	Median	Mean	3rd Qu.	Max
logisic	0.8771	0.912281	0.948276	0.937024	0.964912	0.9824
lasso	0.9824	0.982456	0.982456	0.989474	1	1
svm	0.9824	0.982759	1	0.99304	1	1
Neural_with_LDA	0.9482	0.964912	0.964912	0.972111	0.982456	1

Table 4: Specificity of models

Specificity	Min.	1st Qu.	Median	Mean	3rd Qu.	Max
logisic	0.8235	0.882353	0.970588	0.92352	0.970588	0.9705
lasso	0.8235	0.941176	0.970588	0.93529	0.970588	0.9705
svm	0.9411	0.941176	0.970588	0.95882	0.970588	0.9705
Neural_with_LDA	0.8823	0.941176	0.970588	0.95882	1	1

Table 5: Performance evaluation of models

	Logistic regression	Neural Network with LDA	SV M svm Linea	Lasso Logistic regression
Accuracy	0.9912	0.9646	0.9646	0.9974200
Sensitivity	0.9762	0.9048	0.9048	0.9894737
Specificity	1	1	1	0.93529
Errore class	0.0088	0.0354	0.0354	0.00258

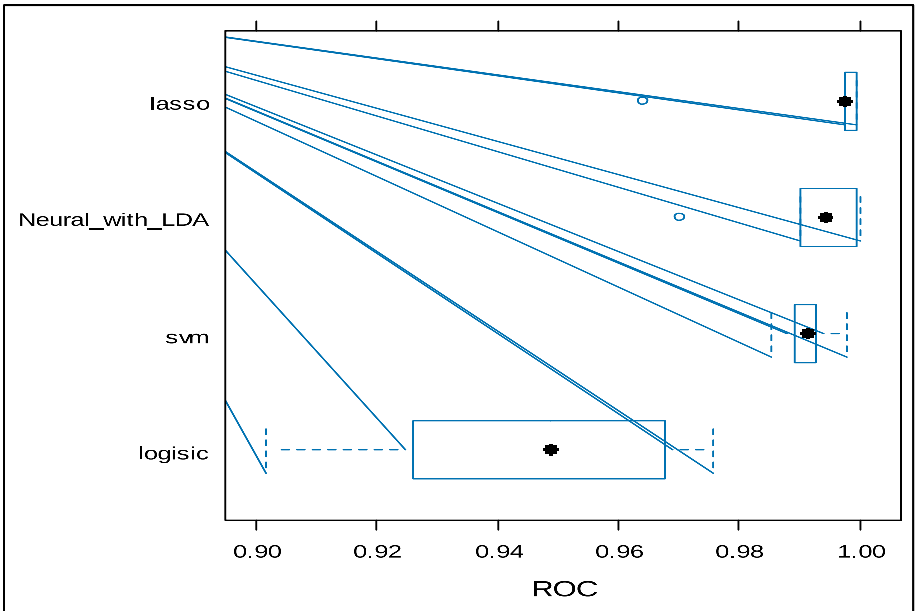

Fig 1: ROC curve comparing the diagnostic performance of the best results obtained for Logistic Regression, SVM, Neural Network, and Lasso.

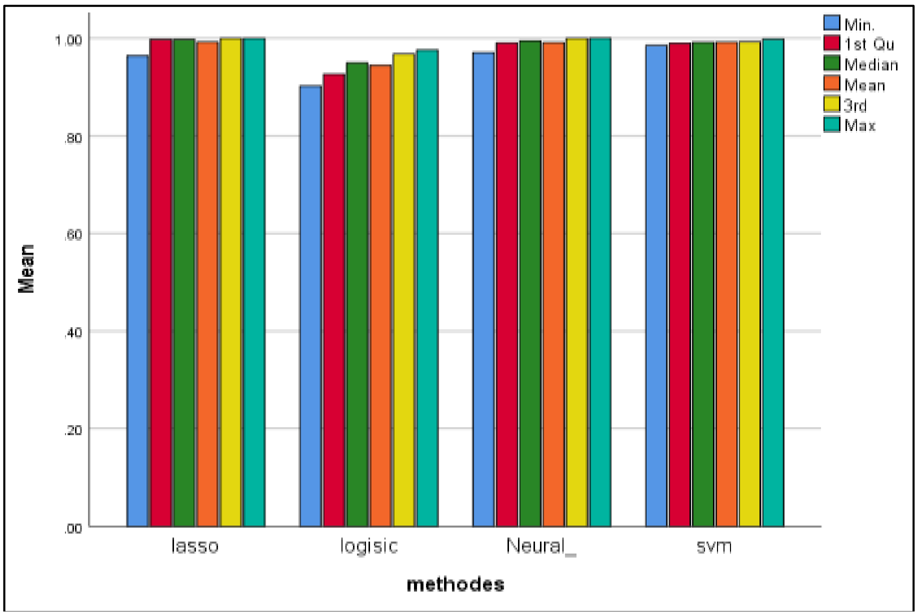

Fig 2: Comparative histogram of the ROC curve obtained for Lasso, Logistic Regression, ANN, and SVM.

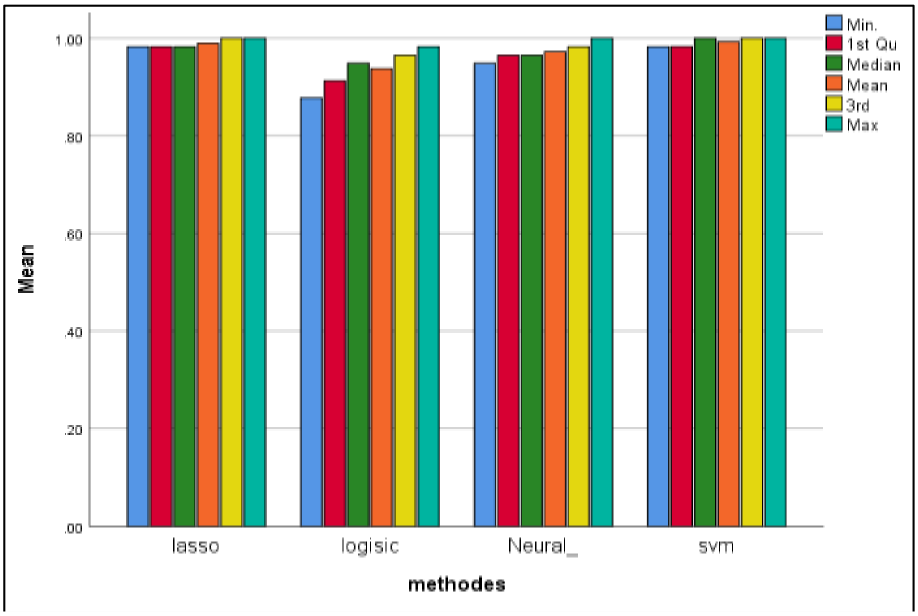

Fig 3: Comparative histogram of the Sensitivity obtained for Lasso, Logistic Regression, ANN, and SVM

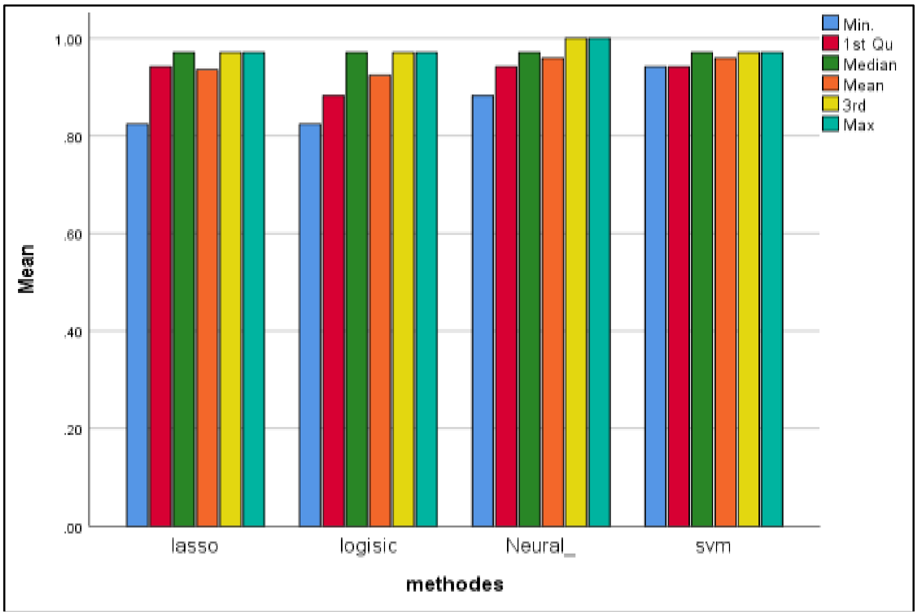

Fig 4: Comparative histogram of the Specificity obtained for Lasso, Logistic Regression, ANN, and SVM.

Classification tasks frequently employ NN, LR, SVM, and Lasso LR. In this study, the four constructed models were compared using the validation dataset after they had been sufficiently trained using the training data to ensure that their output would accurately predict future samples. The results of our study's analysis showed in table (5) that the Lasso LR model has a higher and more accurate ability to diagnose breast cancer to binary logistic regression models, with accuracy of the Lasso LR model getting 99.742%, classification error getting 0.00258 and Sensitivity with (98.947%) and ROC (99.15%). While methods, SVM and NN had lower error classification (0.0354) and accuracy (96.46%) and Sensitivity (90.48 specificity (1%) and ROC (99.1%), The comparison in tables (5) showed that all classification criteria for a SVM & NN model are equal.

4. Conclusion

This research presents a comparison of the diagnostic performance of several machine learning techniques, including logistic regression, Lasso logistic regression, artificial neural networks (NN), and support vector machines (SVM), using the Wisconsin breast cancer data file that was obtained from the UCI Machine Learning Repository or Kaggle, in diagnosing malignant or benign prognosis breast tumors. They were created based on ROC criteria, Sensitivity, Specificity, accuracy, and classification error, and results showed that Lasso Logistic is superior to the logistic regression algorithm for predictive analysis in cancer, outperforming logistic regression for all measures and all models. The test data had the highest error classification accuracy, Sensitivity, Specificity, and ROC values. All classification criteria for an SVM & NN model are equal. The results suggest that this model may provide new opportunities for diagnosing breast tumors because it reduces the problem of multicollinearity and the problem of dimensions by reducing the variables to zero. We also suggest using other regulation methods, such as the Ridge or Elastic regression methods, combining the Lasso and the Ridge. For comparison and categorization, further machine learning methods might be applied.

5. References

1. Osei-Afriyie S, Addae AK, Oppong S, Amu H, Ampofo E, Osei E. Breast cancer awareness, risk factors, and screening practices among future health professionals in Ghana: A cross-sectional study, PLoS One. 2021;l(16):1-17. doi: 10.1371/journal.pone.0253373.
2. Ohno-Machado L, Bialek D. Diagnosing Breast Cancer from FNAs: Variable Relevance in Neural Network and Logistic Regression Models; c1998.
3. Abdolmaleki P, Yarmohammadi M, Gity M. Comparison of logistic regression and neural network models in predicting the outcome of biopsy in breast cancer from MRI findings; c2004.
4. Ayer T, Chhatwal J, Alagoz O, Kahn CE, Woods RW, Burnside ES. Informatics in radiology: Comparison of logistic regression and artificial neural network models in breast cancer risk estimation, *Radiographics*. 2010;30(1):13-22, Jan., doi: 10.1148/rg.301095057.
5. Faradmal J, Soltanian AR, Roshanaei G, Khodabakhshi R, Kasaeian A., Comparison of the performance of log-logistic regression and artificial neural networks for predicting breast cancer relapse, Asian Pacific J Cancer Prev., 2014;15(14):5883-5888, doi: 10.7314/APJCP.2014.15.14.5883.
6. Khosravanian A, Ayat S. Diagnosing Breast Cancer Type by Using Probabilistic Neural Network in Decision Support System, Int. J Knowl. Eng. 2016;2(1):73-76. doi: 10.18178/ijke.2016.2.1.056.
7. Rajbharath R, Sankari L. P.G. Scholar, Predicting Breast Cancer using Random Forest and Logistic Regression; c2017. [Online]. Available: http://ijesc.org/
8. Nourelahi M, Zamani A, Talei A, Tahmasebi S. A model to predict breast cancer survivability using logistic regression, Middle East J Cancer. 2019;10(2):132-138. doi: 10.30476/mejc.2019.78569.
9. Islam MM, Haque MR, Iqbal H, Hasan MM, Hasan M, Kabir MN. Breast Cancer Prediction: A Comparative Study Using Machine Learning Techniques, SN Comput. Sci. 2020, 1(5). doi: 10.1007/s42979-020-00305-w.
10. Özmen-Akyol S. Estimating Breast Cancer Class Using Artificial Neural Network and Logistic Regression Methods, Eskişehir Türk Dünyası Uygul. ve Araştırma Merk. Bilişim Derg. 2022;3(1):26-31. doi: 10.53608/estudambilisim.1052793.
11. Ebrahim M, Sedky AAH, Mesbah S. Accuracy Assessment of Machine Learning Algorithms Used to

Predict Breast Cancer, Data, 2023, 8(2). doi: 10.3390/data8020035.
12. Ibrahim NS, Mohammed NN, Mahmood SW. Multicollinearity in Logistic Regression Model-Subject Review, Iraqi J. Stat. Sci. 2020;17(1):46-53. doi: 10.33899/IQIOSS.2020.165448.
13. Domínguez-Almendros S, Benítez-Parejo N, Gonzalez-Ramirez AR. Logistic regression models, *Allergol. immunopathol.* (Madr). 2011;39(5):295-305. doi: 10.1016/j.aller.2011.05.002.
14. Hooshang N. Advanced Statistics for Testing Assumed Casual Relationships, Multiple Regression Analysis Path Analysis Logistic Regression Analysis,; c2020. [Online]. Available: http://www.springer.com/series/14538
15. Pampel FC. Logistic Regression A Primer SECOND EDITION; c2021.
16. Peng CYJ, Lee KL, Ingersoll GM. An introduction to logistic regression analysis and reporting, J. Educ. Res. 2002;96(1):3-14, doi: 10.1080/00220670209598786.
17. Herawati N, Herawati N, Nisa K. Selecting the Method to Overcome Partia l and Full Multicollinearity in Binary Logist is Model.pdf AUTHOR 15% Overall Similarity Excluded from Similarity Report Selecting the Method to Overcome Partial and Full Multicollinearity in Binary Logistic Model, Int. J. Stat. Appl. 2020, 2022;3:55-59. doi: 10.5923/j.statistics.20201003.01.
18. Meier L, Geer VDS, Bühlmann P. The group lasso for logistic regression; c2008.
19. Melkumova LE, Shatskikh SY. Comparing Ridge and LASSO estimators for data analysis, in Procedia Engineering, Elsevier Ltd; c2017. p. 746-755. doi: 10.1016/j.proeng.2017.09.615.
20. Shanmuganathan S. Studies in Computational Intelligence 628 Artificial Neural Network Modelling; c2016. [Online]. Available: http://www.springer.com/series/7092
21. Goode P. Neural Network Control of Nonlinear Discrete-Time Systems; c2006.
22. Hu YH, Hwang JN. Handbook of neural network signal processing. CRC Press; c2002.
23. Yadav N, Yadav A, Kumar M. springer briefs in applied sciences and technology. Computational intelligence: An Introduction to Neural Network Methods for Differential Equations. c2015. [Online]. Available: http://www.springer.com/series/10618
24. Jeswal CSKS., World Scientific Solving Algebraic Equations; c2021.
25. Livshin I. Artificial Neural Networks with Java. Apress; c2022. doi: 10.1007/978-1-4842-7368-5.
26. Zeng Z, Wang J. Advances in Neural Network Research and Applications (Lecture Notes in Electrical Engineering, 67); c2010.
27. Rezaeian A, Rezaeian M, Khatami SF, Khorashadizadeh F, Moghaddam FP. Prediction of mortality of premature neonates using neural network and logistic regression, *J.* Ambient Intell. Humans. Comput. 2022;13(3):1269-1277. doi: 10.1007/s12652-020-02562-2.
28. Park HA. An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain, J Korean Acad. Nurs. 2013;43(2):154-164. doi: 10.4040/jkan.2013.43.2.154.