JQCM

University Of AL-Qadisiyah

Investigating the Applicability of Logistic Regression and Artificial Neural Networks in Predicting Breast Cancer

Yousra Abdulaziz Mohammed[a],   Eman Gadban Saleh[b]

[a]College of Medicine, University of Fallujah, Anbar, Iraq,
Email: yosra_mohammed@uofallujah.edu.iq
[b]College of medicine, University of Fallujah, Anbar, Iraq,
Email: emanghadbansaleh@uofallujah.edu.iq

ARTICLE INFO

ABSTRACT

Breast cancer has become recently the most common cancer and a major cause of death among women all over the world and especially in developing countries like Iraq. This study aims to predict the type of breast tumor whether benign or malignant through the different models that were built using logistic regression and neural networks which is expected to be helpful for oncologists in diagnosing the type of breast tumor. Four models were set using binary logistic regression and two different types of artificial neural networks namely multilayer perceptron (MLP) and radial basis function (RBF). Both validated and trained models were evaluated using different performance metrics like accuracy or correct classification rate (CCR), receiver operating characteristic (ROC) curves, area under ROC curve (AUC), sensitivity and specificity. Dataset has been downloaded from the machine learning repository of University of California, Irvin (UCI ml repository) that consists of 9 attributes and 699 valid instances.

Firstly, some preprocessing was done to cleanse the data, then the models were built using the Logistic Regression method and Artificial Neural Networks and a comparison was done to find out which model will give the highest performance. Each model was validated with a different dataset than that used for developing the models. The analysis of the results showed that the Radial Basis Function neural network model is the best classifier in the prediction of the type of breast tumors since it had recorded the highest performance in terms of correct classification rate (accuracy), sensitivity, specificity, and AUC among all other models.

MSC:

1 . Introduction

Breast cancer is one of the most common types of cancer in the world, especially in the developing countries like Iraq [1, 2]. It is accounting for 25% of all malignancy, with about new 1.57 million cases in 2012[1]. It is the first – ranked cause of woman cancer-related deaths. Although fundamental

Corresponding author: Yousra abdulaziz Mohammed

Email address: yosra_mohammed@uofallujah.edu.iq

Communicated by : Alaa Hussein Hamadi

improvement in the survival for this disease has been recorded in countries with high-resource, the risk go on to rise, recording high mortality rates in developing countries[1]. The Iraqi cancer Registry disclosed the breast cancer is the highest rates of cancer cases (19.1%) and the highest annual estimate of cancer in women (25.8 per hundred thousand of the female population) .The second – ranked incidence of cancer mortality was breast cancer (2.7/100,000 populations)[3]. According to a study, survival rate is 88% after 5 years of diagnosis and 80% after 10 years of diagnosis in early stage which means that about 88% of women diagnosed with breast cancer will survive for at least 5 years after their early diagnosis, therefore it is necessary to detect breast cancer at earliest stage possible[4].

 Machine learning has become more influential in diagnosing cancer, because It allows deductions or inference to be made that classical statistical procedures could not make[5].

In order to assist oncologists making the right diagnosis of biopsy in breast cancer, a classification model known as a classifier can be very helpful. The Classification problem refers to predicting the target class of new observations, from a given set of predictive variables from the population dataset.

Since the outcome of biopsy can confirm the existence/ absence of the malignancy it is hence considered a binary outcome.

 Logistic Regression (LR) have been applied increasingly in many fields particularly the medical fields, and is a perfect statistical algorithm for binary classification, that is evaluating the correlation between one or more categorical or continuous predictor variables and a dichotomous dependent variable[6]. Logistic regression technique has the ability of assigning distinct datasets to predefined classes, the distinction is done by setting up the discrimination rules, these rules are estimated through the training phase and can be used to assign the new observations into the classes defined formerly[7].

 Methods of variable selection differ according to the problem. It is essential to include all relevant variables in the model. Some researchers propose inclusion of all clinical and other predictive variables in the model regardless of their significance to get a better model fit to the data. Yet, more variables will affect the coefficient in the model and lead to over-fitting model. Besides, a model with many insignificant predictors will produce less classification accuracy and it would be hard to explain the results. Commonly, statistical model building techniques attempt to minimize the number of variables to get a numerically stable and generalized model, but this can cause in a large standard errors. Variables selection can be done in two ways filter and statistical[8, 9]. For filter method, the variables are reduced according to their importance as was done in similar research. On the other hand, statistical method for variable selection can be done by either of the following methods[10];

- "Enter: A procedure for variable selection in which all variables in a block are entered in a single step.
- Forward Selection (Conditional).
- Forward Selection (Likelihood Ratio).
- Forward Selection (Wald).
- Backward Elimination (Conditional).
- Backward Elimination (Likelihood Ratio).
- Backward Elimination (Wald)."

In forward selection, the significant effects once entered then could not be removed from the model. For backward elimination, removed effects from the model, cannot be entered again. While, for stepwise selection which are the method of focus in our study the variables already included in the model do not need to remain, they can be entered into or eliminated from the model in a certain manner that every step of forward selection could be followed by a backward elimination step or more. The stepwise selection procedure stops if no additional effect is added to the model[11]. In this study, two different variable selection procedures were implemented, namely Enter and Stepwise methods to establish logistic regression models.

Artificial neural networks (ANN) are commonly used as a robust decision making systems especially for medical diagnosing after being trained using historical data set.   ANNs advantages can be summarized in that tuning neural weights is done online with no need to any pre-training phase, and persistence and performance systems is ensured. ANN is a powerful classifier that represents a nonlinear relationship between input and output. Basically, a simple ANN consists of three layers, an input layer, hidden layer/s and an output layer. At the input layer the inputs are weighted, i.e. each input value is multiplied by certain weight. At the hidden layer, all weighted inputs along with a bias are summed. Finally at the output layer the summed value obtained   is converted to activation signal using transform function. The ANN is trained with a learning algorithm according to the type of the given problem. Generally the learning algorithms are either supervised learning, unsupervised learning or reinforcement learning [12].

This study aims to evaluate the performance of   two techniques, logistic regression and neural networks in order to determine which of the used methods is more powerful in classifying the type of breast tumor in benign or malignant classes.

2. Material and Methods
2.1 Data

The dataset used in this study, is the breast-cancer-Wisconsin.data file which was collected from UCI machine learning repository[13].The dataset consists of observations of 699 breast fine needle aspirations (FNA)s. It is arranged in 11 columns each row represents observations belong to a patient's breast FNA that was obtained from medical analysis. The first column is an identification code associated with each patient; the following 9 columns are the features used to analyze each FNA obtained from patient breast tumor; clamp thickness (range to which cell aggregates , mono- or multilayered), uniformity of cell sizes, uniformity of cell forms, marginal adhesion (coherence of the marginal cells of the cell aggregates), size of the single epithelial cell(diameter of the inhabitance of the biggest cells comparative to erythrocytes), Bare nuclei (the ratio of single cell nuclei that were freed from encirclement cytoplasm), chromatin blandness, nucleolus normality, and mitosis[14]. The last column is the dependant variable (cancer type; 4 for malign and 2 for benign tumors).

 "All malignant aspirates were histologically confirmed whereas FNAs diagnosed as benign masses were biopsied only at the patient's request. The remainder of benign cytologies were confirmed by clinical reexamination 3 and 12 months after the aspiration. Masses that produced unsatisfactory or suspicious FNAs were surgically biopsied"[15].

All the independent variables have numerical values ranging from 1 to10, and these values were obtained through medical analysis or lab tests. The distribution of the dependant variable Class is; Benign: 458 (65.5%) and Malignant: 241 (34.5%).

The first step in this study was converting the data into an Excel sheet to make it easier to build the statistical model which is generated and analyzed using SPSS, V19.0, SPSS Inc. using LR algorithm, then the data was imported to (SPSS) program for processing.

The second step is data cleaning; missing values is a well known issue that exists in datasets. There are several methods to overcome missing values like Listwise or case deletion, substituting the missing values with mean or mode of that variable and other methods[16]. For our study the missing values were replaced by the mean of the nearby attribute values. As it is not required in the design and analysis of our model, the identifier number (Id)column was removed.

2.2 Logistic Regression (LR) Model

The conditional probability for dependent variable to occur is given by the logistic function[17],

$$Probability(Y = 1) = \frac{e^{-z}}{1+e^{-z}} \tag{1}$$

Where probability estimates are between 0 and 1 because of the logistic transformation, z is also called logit. The logit is a linear multiple regression model of the independent variables

$$z = \acute{\beta} \acute{X} = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \cdots \beta_n * X_n \tag{2}$$

Where $\beta_0 \ldots \beta_n$ are coefficients of the independent variables calculated by estimation of the maximum likelihood, $X_1 \ldots X_n$ are independent variables and n is the number of explanatory variables.

While reference probability is defined as,

$$p(y = 0) = 1/(1 + e^{-z}) \tag{3}$$

the log(odds), or log-odds ratio, is defined by,

$$log(odds) = ln\left[\frac{p}{1-p}\right] = z \tag{4}$$

and expresses the natural logarithm of the ratio between the probability that an event will occur, *p(Y=1)*, to the probability that it will not occur *p(Y=0)*, it is found by calculating the probability of each event. Odds ratio measure the incidence when the independent variable increases by one unit. The odds ratio is defined as,

$$\frac{p}{(1-p)} = e^{\acute{\beta} \acute{X}} \tag{5}$$

For the first method two models were set, a full model using the standard ENTER method with all 9 attributes and a reduced model using the stepwise forward selection (Wald) method. Stepwise selection method tests the entry of variables according to the significance of the score statistic, while removal testing is done according to the probability of the Wald statistic[10], the model was developed

with only 5 attributes which   are tumor thickness, uniformity of cell size, marginal adhesion, bare nuclei, and bland chromatin since they were statistically significant at the level of 0.05 using Wald statistic.

Validation of the model is very important to measure the stability and robustness of the coefficients resulting from logistic regression and a crucial part of the process of model-building[18]. The validation is using different data set pertaining the values of the coefficient as for the training data to calculate the percentage of correct classifications. The percentage of correctly predicted samples from the training samples must be ≥ to the validated samples[19].

Many statistical tools for model performance validation in binary logistic regression are available like data splitting, repeated data-splitting, jackknife technique and bootstrapping [20]. For this purpose, the data-splitting technique was used in our study, where the data had been randomly divided into two groups; the first consisting of 80% of the data (550) sample was used for developing the LR model with 373 benign and 177 malignant, and the second group consisting of 20% of data (149) sample (85 benign, 64 malignant)was used for validating the two models.

 The training data was used at first to fit both full and reduced models then we apply the validation data to the fitted models to evaluate the model's performance. The obtained posterior probability for malignant class was considered and its value was then classified into two categories; posterior probability in range of (0-0.5)= benign, and posterior probability in range of (0.5-1)= malignant. Results obtained are then evaluated in terms of measures such as ACC, Specificity, Sensitivity, and ROC curve area.

2.3 Neural Networks

In our study, two types of ANN were used. The first one is Multilayer Perceptron (MLP) network (Fig.1) which is a well known network architecture that has been used in medical, engineering, mathematical modeling research. In MLP, a fixed value (bias) along with weighted sum of   inputs are propagated to the hidden layer via a transfer function to generate the output, and the   topology of feed-forward layers arrangement of units is called Feed Forward Neural Network (FFNN). The learning ability of the MLP is highly increased by the hidden layer.   The input is modified by the activation function of the network so as to give a required output.   Model building is strongly affected by the hidden nodes number, hidden layers number, and the type of activation function selection[21]. The output of a MLP NN is given below:

$$y(j) = T \left(\sum_{k=0}^{n} w_k(j) * x_k(j) + c \right) \tag{6}$$

Where $y(j)$ is the output value, x_k is the input vector, T is the transform function, c is a constant, w_k is the vector of weights, n is the size of input vector. The equation is in discrete time j [12] .
The second type of NN used is radial basis function neural network RBF which is based on supervised learning. RBF NN are efficient in modeling nonlinear data and training this type of NN can be done in one stage counter to MLP. In the hidden layer RBFNN uses nonlinear Gaussian transfer function whereas in the output layer it uses a linear summation transfer function. The real values of the n-dimensional input vector X is fed to all units in the hidden layer at the same time (Fig.2). The Gaussian RBF is given by;

$$\emptyset(\|x - x^{(i)}\|) = exp\left(-\frac{(\|x-x^{(i)}\|)}{2\sigma_i^2}\right) \tag{7}$$

Where the functions $\emptyset(\|x - x^{(i)}\|)$, *i=1,2,...,N* are called the RBFs, where a p-norm (often the Euclidean 2-norm) denotes $\| . \|$, $x^{(i)}$ is the basis function centre and σ_i is its radius. A linear combination of basis functions can be used for approximation of a nonlinear function. The output : $R^n \rightarrow R$, of the network is thus,

$$F(x) = \sum_{i=1}^{N} w_i \emptyset(\|x - x^{(i)}\| \tag{8}$$

where N is the number of neurons in the hidden layer and the real parameters $w_i, i = 1, 2, \ldots, N$ are the weights of the linear output neurons[22].
To train RBF networks, once the type of radial basis function is selected, all needed to do is choosing the functions' dimensions and centers and estimating the output neuron weights.
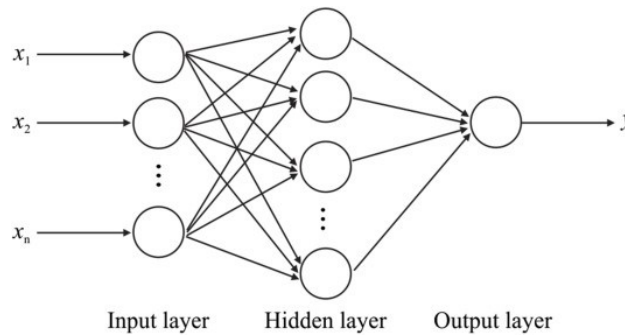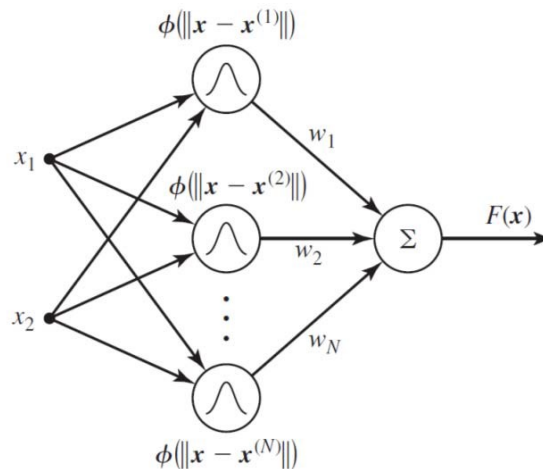

Figure 1. Multilayer Feedforward MLP

Figure 2. Radial Basis Function NN

For the ANN, two models were developed using two different types of NN, namely MLP and RBF. The architecture of the MLP neural network had four layers; the input layer consisted of 9 input elements, corresponded to the data taken from cytology, then two hidden layers with sigmoid activation function, the first one had 7 nodes while the second hidden layer consisted of 5 nodes and the output layer with 2 neurons, representing 0 for benign and 1 for malignant lesions. A back propagation algorithm based on scaled conjugate optimization technique was used to model MLP for our dataset. To get the optimum

neural network structure, a considerable number of neural networks has been simulated by changing the number of hidden layers, hidden nodes, iterations and learning rates. Whereas the feed forward topology of RBF network developed for this work was composed of 3 layers, input layer with the 9 input elements, a single hidden layer with a nonlinear RBF activation function and 9 neurons fully interconnected to the output layer units and a linear output layer with 2 elements. The error function E used to index the learning efficiency of both neural networks was the Sum of Squared Error (SSE) criterion function which had to be minimized over the given training set. The performance of the NN models was determined by dividing the dataset into two separate sets 70% of samples for the training and 30% for the validation.

After the networks had been trained perfectly using the training data, each network was tested by presenting the testing set to the trained network and a diagnostic output vector of 0's and 1's was generated.

2.4 Performance Metrics

  Accuracy which is the percentage of correct predictions is the most used measure in classification task. Sensitivity and specificity have to be calculated because the first indicates the performance of classification for minority class, while the second indicates the proportion of majority samples that are correctly identified. Also the area under a ROC curve (AUC) was used to evaluate the performance of the feature selection method[23]. For our work, the two models were evaluated using these metrics (Equations 9-13) based on the confusion matrix shown in Table 1.

$$CCA = (TP + TN)/(TP + FP + TN + TP) \qquad (9)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \qquad (10)$$

$$\text{Specificity} = TN/(TN + FP) \qquad (11)$$

$$AUC = \frac{1+TP_{rate}-FP_{rate}}{2} \qquad (12)$$

Where, TP rate is sensitivity and

$$FP_{rate} = \frac{FP}{FP+TN} \qquad (13)$$

Table 1. Confusion Matrix

predicted class → Actual class ↓	Benign (0)	Malignant (1)
Benign (0)	TN	FP
Malignant (1)	FN	TP

3. Results

In this study, different models were set using IBM SPSS statistics 19 software and the performance of the classifiers was compared. The dataset related to breast cancer was downloaded from UCI-Machine Learning repository, and was fed to our Logistic-regression models, MLP, and RBF neural networks. Each classifier was well trained with the dataset   and a Model is set and validated with test samples, then results were obtained. The results of training LR full model using the training sample is shown in

Table 2 where the values of logistic regression parameters, standard errors, Wald statistics and p-values of the logistic regression model are computed. Considering all available variables, the logit of the full model is given by,

Logit1= - 10.112 +.586*clump thickness + 0.218* uniformity of cell size + 0.159*uniformity of cell shape + 0.304*marginal adhesion - 0.110*single epithelial cell size + 0.394*bare nuclei+0.483*bland chromatin + 0.115*normal nucleoli + 0.487*mitosis.

From Table 2, we find that small p-values of clump thickness, marginal adhesion, bare nuclei, and bland chromatin indicate that they are most significant predictor of malignancy in the model at level of 0.05. Also, the coefficients of the reduced model were computed from training the model by the stepwise method using the training sample is shown in Table 3. from which the logit of the reduced model is given by,

Logit 2(stepwise)= -9.911+0.627* clump thickness +0.376* uniformity of cell size +0.282* marginal adhesion +0.396* bare nuclei +0.566* bland chromatin.

Table 2. Parameter Estimations of the Full LR model fitted to the training sample.

Variable	Coefficient estimate	Standard error	Wald	Sig.
clump_thickness	.586	.180	10.648	.001*
uniformity_of_cell_size	.218	.238	.838	.360
uniformity_of_cell_shape	.159	.260	.374	.541
marginal_adhesion	.304	.130	5.476	.019*
single_epithilial_cell_size	-.110	.243	.205	.650
bare_nuclei	.394	.105	13.963	.000*
bland_chromatin	.483	.204	5.589	.018*
normal_nucleoli	.115	.137	.708	.400
metosis	.487	.477	1.042	.307
Constant	-10.112	1.460	47.973	.000
*significant at level of 0.05				

For the neural network, both MLP and RBF models has shown a considerable improvement in all performance metrics than those scored by logistic regression models, and RBF NN has preceded all the other models developed in this work with highest correct classification rate of 95.4%, sensitivity of 98.5%, AUC of 96.125%, except for specificity which was 93% the same for both MLP and RBF. Table 4 and Fig.3 show a comparison of the performance of the logistic regression models and two types of neural networks on testing samples in terms of percentage of correct classification rate (accuracy), sensitivity, specificity, and areas under receiver operating characteristic curve AUC.

Table 3. Parameter Estimations of the reduced LR model (stepwise) fitted to the training samples.

Variable	Coefficient estimate	Standard error	Wald	Sig.
clump_thickness	.627	.172	13.323	.000
uniformity_of_cell_size	.376	.149	6.409	.011
marginal_adhesion	.282	.122	5.320	.021
bare_nuclei	.396	.099	16.021	.000
bland_chromatin	.566	.193	8.643	.003
Constant	-9.911	1.312	57.064	.000
*significant at level of 0.05				

Table 4. Comparative performance of the four models on validation samples

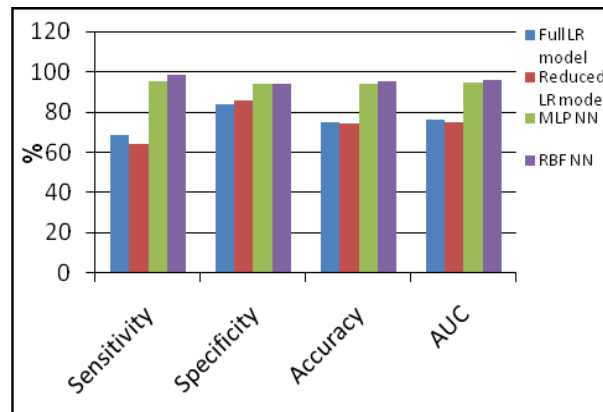| Model	Full LR model	Reduced LR model	MLP NN	RBF NN
Sensitivity %	68.75	64.1	95.1	98.5
Specificity %	83.53	85.88	93.8	93.8
Accuracy (CCR) %	74.7	74.03	94.2	95.4
AUC %	76.1	74.97	94.45	96.125

Figure 3. Comparison of the average sensitivity, specificity, accuracy, and AUC obtained for the full, reduced LR models, MLP, and RBF artificial neural networks

4. Discussion

ANN and LR are widely used for tasks of prediction or classification. In this work, comparison of the four models developed was based on the validation dataset after the models had been sufficiently trained with the training data to assure whether the output of the these models will predict future samples precisely. The first logistic regression model set was full model which included all the nine covariates and the second was a reduced model. The second model built with logistic regression was the reduced model using the stepwise method where the variables with the largest Wald test p-value has been removed which were; uniformity of cell shape, single epithelial cell size, mitosis, and normal nucleoli, retaining the coefficients of the significant covariates only, but the output of the reduced model did not show any improvement in any of the metrics used except for specificity as shown in Table 4. & Fig.3.

The result analysis of our study showed that the ability of RBF NN to diagnose breast cancer is superior to Binary Logistic Regression models (both full and reduced) and to MLP achieving highest and most accurate results where the accuracy of the RBF NN model was 95.4% and a sensitivity of 98.5% followed by MLP. While logistic regression method showed much less accuracy (74.7% for full model and 74.03% for reduced), sensitivity (68.75% for full model and 64.1% for reduced), specificity (83.53% for full model and 85.33% for reduced), and AUC (76.1% for full model and 74.79% for reduced), in comparison with neural networks.

5. Conclusion

This paper represents a comparative study of the diagnosing performance of two different machine learning techniques namely logistic regression and artificial neural networks in the prognosis of breast tumors weather malignant or benign using the breast-cancer-Wisconsin.data file which was collected

from UCI machine learning repository. The diagnosing performance of two types of artificial neural networks and binary logistic regression through the different models that were built, was compared based on sensitivity, specificity, AUC, and accuracy or CCR criteria. The results showed that using ANN in predictive analysis in oncology is more powerful than logistic regression algorithm, RBF outperforms MLP and logistic regression for all metrics. The sensitivity CCR, AUC values for RBF on testing data were the highest. The findings indicate the possibility to open new opportunities in the diagnosis of breast tumors.

References

[1] NA. Alwan "Breast cancer among Iraqi women: Preliminary findings from a regional comparative Breast Cancer Research Project," *Journal of global oncology*, vol. 2, issue 5, pp. 255, 2016.

[2] World population review/Developing countries 2020, available at:
 https://worldpopulationreview.com/countries/developing-countries/ extracted on June 19, 2020.

[3] Iraqi Cancer Board. Results of the Iraqi Cancer Registry Baghdad, Iraq, Iraqi Cancer RegistryCenter, Ministry of Health, 2015. available at: https://moh.gov.iq/upload/upfile/ar/833.pdf   accessed on January 27 2020.

[4] M. Rathi, A.K. Singh, "Breast Cancer Prediction using Naïve Bayes Classifier," International Journal of Information Technology & Systems, vol.1, issue 2, pp.77-80, 2012.

[5] JA Cruz, DS Wishart, "Applications of Machine Learning in Cancer Prediction and Prognosis," Cancer Informatics. vol.2, pp. 2-21. 2006.

[6] https://science-network.tv/logistic-regression/. Accessed   September   24,   2019.   Ronny Gunnarsson and first published June 22, 2014. Last revised August 30, 2019.

[7] DW Hosmer, S. Lemeshow Applied logistic regression. New York: Wiley; 1989.

[8] PC Austin, JV. Tu, "Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality," Journal of Clinical Epidemiology. vol.57, issue 11, pp. 1138-1146, 2004.

[9] R. Genuer, JM Poggi, C. Tuleau-Malot, "Variable selection using random forests," Pattern Recognition Letters. vol.31, issue 14, pp. 2225-2236. 2010. DOI:10.1016/j.patrec.2010.03.014

[10] IBM SPSS Regression 22. Copyright IBM Corporation 1989, 2013.

[11] Z Bursac , CH Gauss , DK Williams , DW Hosmer, "Purposeful selection of variables in logistic regression," Source Code for Biology and Medicine, vol.3, issue 17, 2008. doi:10.1186/1751-0473-3-17

[12] M.G Kanojia,   S. Abraham, 2016, December. "Breast cancer detection using RBF neural network," in 2nd International Conference on Contemporary Computing and Informatics (IC3I) 2016, pp. 363-368, IEEE.

[13] D Dua, C Graff. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science

[14] F Leisch,   E. Dimitriadou, (2015). Machine learning benchmark problems. Viewed on 24.09.2019 from https://cran.r-project.org/web/packages/mlbench/mlbench.pdf.

[15] WH Wolberg, OL Mangasarian, "Multisurface method of pattern separation for medical diagnosis applied to breast cytology," in Proceedings of the national academy of sciences, United States of America, Dec, 1990,   pp.9193-6.

[16] H. Kang, "The prevention and handling of the missing data," Korean journal of anesthesiology," vol.64, issue 5, pp. 402–406, 2013. doi:10.4097/kjae.2013.64.5.402

[17] J Schwarz, p.Heidi Bruderer Enzler, (2014). Research Methodology: Tools; Applied Data Analysis (with SPSS). Lecture 08: Logistic Regression Analysis, Lurcern University. Available at, http://www.schwarzpartners.ch/Applied_Data_Analysis/Lecture%2008_EN_2014%20Logistic%2 0Regression%20Analysis.pdf   extracted on September 30, 2019.

[18] J. Shao, "Linear Model Selection by Cross-Validation,"   Journal of the American Statistical Association, vol.80, No.422, pp. 486-494, 1993.

[19] H. Yusuff, N. Mohamad, U.K. Ngah, A. Yahaya, "Breast cancer analysis using logistic regression," International Journal of Research and Reviews in Applied Sciences, vol.10, issue1, pp.14-22, 2012.

[20] S. Rana, H Midi, SK Sarkar, "Validation and Performance Analysis of Binary Logistic Regression Model," in Proceedings of the WSEAS International Conference on Environment, medicine and health sciences; Penang, Malaysia, 2010, pp.51-55.

[21] J. Padmavati, "A comparative study on breast cancer prediction using RBF and MLP," International Journal of Scientific & Engineering Research, vol.2, issue1, pp. 1-5, 2011.

[22] H. Jafarnejadsani, J.Pieper, J. Ehlers, "Adaptive control of a variable-speed variable-pitch wind turbine using radial-basis function neural network," IEEE transactions on control systems technology, vol.21, issue 6, pp. 2264-2272, 2013. DOI: 10.1109/TCST.2012.2237518

[23] S Miri Rostami, M Ahmadzadeh, "Extracting Predictor Variables to Construct Breast Cancer Survivability Model with Class Imbalance Problem," Journal of AI and Data Mining, vol.6, issue2, pp.263-276, 2018.   DOI: 10.22044/JADM.2017.5061.1609