

Using Multilayer Perceptron and Deep Neural Networks for the Diagnosis of Breast Cancer Classification

¹Igodan, C. Efosa, ²Ukaoha, Kingsley, C.

^{1,2}Department of Computer Science, University of Benin, Benin City, Nigeria
Email: charles.igodan@uniben.edu, kingsley.ukaoha@uniben.edu

Article Info

Keywords:

Breast Cancer, Multilayer Perception, Back Propagation, Classifier Models

Received 29 January 2020

Revised 17 March 2020

Accepted 06 April 2020

Available online 1 June 2020



<https://doi.org/10.37933/nipes/2.2.2020.3>

<https://nipesjournals.org.ng>
ISSN-2682-5821/© 2020 NIPES Pub.
All rights reserved.

Abstract

The very long delay that is suffered by patients of breast cancer in their early stages in low-income countries is due to access barriers and quality deficiencies in the care of cancer giving rise to the need for an alternative and efficient computer-based diagnostic system for the early detection and prevention of the disease. The early detection and improved therapy still remain a crucial approach for the prevention and cure of breast cancer. To this end, recent research looks into the development of different classifier models for the classification of breast cancer. This paper investigates the potentials of applying multiple neural network architectures with increased number of hidden layers and hidden units. The network architectures have one-hidden-layer, two-hidden-layer and three hidden layer (deep neural network) architectures respectively using the backpropagation training algorithm for the training of the models. The experimental results show that by applying this approach the models yield efficient and promising results

1. Introduction

Breast cancer disease is one of the leading causes of death among women the world over [1], [2],[3],[4] and [5]. Most breast cancer cases occur in women aged 40 and above but certain women with high risk characteristics may develop breast cancer at a younger age. Cancer is a disease in which cells become abnormal and form more cells in an uncontrolled way. With breast cancer, the cancer begins in the tissues that make up the breasts. The cancer cells may form a mass called a tumor. That may also invade nearby tissue and spread to lymph nodes and other parts of the body. The most common types of breast cancer are the Ductal carcinoma and Lobular carcinoma. Ductal carcinoma cancer begins in the ducts and grows into surrounding tissues. About 8 in 10 breast cancers are this type. Lobular carcinoma cancer begins in lobules and grows into surrounding tissues. About 1 in 10 breast cancers are of this type [4]. Identifying the breast cancer tumor quickly and accurately, either benign (non-cancerous) or malignant (cancerous), is very critical for taking the correct and right treatment medically. It is very difficult to describe the main morphological features of breast cancer owing to its complex

nonlinear relationship using common traditional linear regression methods. In [3], among a great variety of classification techniques suggested so far for the medical diagnosis, neural networks have been one of the most popular methods that consistently demonstrated its strength and potentials in solving practical classification problems. Most of the research carried out on the medical diagnosis of breast cancer was done using the Wisconsin breast cancer database (WBCD) in neural network literature [1], [3],[6], and [7]. This paper adopted the use of a feed forward multilayer neural network with varying sizes and depths in terms of the number of layers and neurons. The objective of the proposed models is to determine the best neural network architecture best for the classification of the breast cancer irrespective of the depth and size. Through simulation, the models were experimentally compared using various metrics as the ROC, confusion metrics and so on to show the one that obtains the highest classification accuracy.

1.1. Related Work

Recently, different methods proposed for the detection and classification of breast cancer into benign or malignant cases can either fall under be the statistical methods, neural network, evolutionary computation methods and / or combination of any of the methods. However, using the neural network for breast cancer diagnosis has recently received a good deal of attention [3], and [6] as this is due to its ability to represent the behaviour of linear or nonlinear function multidimensional and complex [6]. Among other methodologies, the classification model is good at the detection and classification of breast cancer data [5]. Some of the various simple, hybrid or adaptive artificial techniques used are reviewed in this literature. [2] present three diagnosis systems using pattern recognition based on genetic algorithm and neural networks. The system performances are estimated by classification accuracy and compared with similar methods without feature selection. The paper concluded that the results of the hybrid methods (GA-GRNN, GA-RBF, and GA-RBRF) gives better accuracy than the simple methods (GRNN, RBF, and RBEF) and can be helpful for physicians of their patients' diseases. [8] built an artificial neural network model for detection of breast cancer based on Image registration techniques. The performance of the systems is analyzed on the basis of mean squared error for different number of neurons of ANN computed on hit and trial method showing improved classification efficiencies. In [9], a genetic-neuro system classification is proposed for the classification of breast cancer. The study proved that the experimental results obtained showed that the classification model performs better than the conventional neural networks. [10], presents an evolutionary fuzzy ARTMap algorithm on breast cancer dataset. The purpose of the study is to present the strength of fuzzy artmap using GA to optimize its parameters for improved classification performance in accuracy over the ordinary neural networks. In the paper, the main conclusion of the solution is that it could be applied on any problem, give high accurate performance result and solve drawback of user tuning for fuzzy artmap parameters. In [1], an investigation was carried out on the potential of applying the feed forward neural network architecture for the classification of breast cancer using back-propagation training algorithm. The

paper presented the result of a comparison among ten different hidden neurons initialization methods. In the study, the network architecture with six hidden neurons gave the highest diagnostic performance of 99.28% accuracy. [6] proposed a new approach based on the concept of feed forward neural networks and Island differential evolution propagation algorithms to train the network on breast cancer classification problem. In the study, the Island differential evolution neural network approach driven by the learning algorithm works well in terms of accuracy efficiency and reliability. In [7], a parallel approach by using feed forward network techniques and back propagation learning algorithm is proposed to help in the diagnosis of breast cancer. The performance of the network is evaluated and experimental result reveals that by applying parallel approach in neural network model, efficiency is achieved. A genetic neural network classification model is developed by [3]. The proposed model optimizes the weights and threshold and also reduces the size of the network by identifying the feature subset using GA. In the study, simulation results show that the developed model achieved dimensional reduction and improved classification accuracy and excellent efficiency. [11] presented a novel hybrid intelligent method for detection of breast cancer. The proposed method includes two main modules: clustering module and the classifier model. In the study, the best classifier with the highest rate of accuracy is chosen in order to recognize the breast cancer. Simulation shows that the best classifier obtain 99.11% accuracy rate. In [12], a fast learning neuro-evolutionary technique that evaluates artificial neural networks using Cartesian Genetic Programming (GA-PANN) to detect the presence of breast cancer is proposed. The developed system produces fast and accurate results when compared to contemporary work done in the field. In the study, the error of the model comes out to be as low as 1% for type I (classifying benign samples falsely as malignant or classifying true instance as false) and 0.5% for type II (classifying malignant sample falsely as benign or classifying false instance as true). [13] introduces four new methods for extracting the speculation features of a detected breast lesion on mammography by segmenting the contour of the lesion in a number of regions which are separately analyzed, determining a characterizing speculation feature set using neural network. In the paper, the performance of the methods is analyzed depending on the number of regions in which the contour is segmented and the performance related conclusions are stated for each of the methods. [14] proposed an evolutionary neural network pruning method for breast cancer diagnosis factors elimination problem. The GAs was used for pruning neural network structure and the investigation of the most appropriate subset of input parameters of ANNs that can provide reliable medical diagnosis. In the paper, the findings indicate that there is high level of redundancy in the original full-sized breast cancer diagnosis data. In [15] a novel approach to ANN topology optimization using evolutionary algorithms for breast cancer classification problems was presented. In the paper, the proposed solution proved to be able to reach a good level of optimization pruning the original architecture, (returned by the fixed topology optimization approach) to a solution, leaving unchanged the accuracy level of the system. [16], proposed a probabilistic neural network (PNN) to devise a decision support system (DSS) to diagnose the type of breast cancer in patients. In the work, the proposed model obtained high

performance with a sensitivity of 1, specificity of 0.98% and accuracy of 0.99 respectively. [17] presented a computerized breast cancer diagnosis prototype with GAs and neural network to reduce the time taken and indirectly reducing the probability of death. The research work through simulating the training process with increasing number of hidden layers and hidden neurons to identifying the best solutions shows that architecture with 9-3-1 (one-hidden layer) is still the best architecture for the dataset. However, the work shows that there are still works to do as to comparing the results with existing classifier models as well as refining the interface. An adaptive combination of genetic algorithm and ART neural network for breast cancer diagnosis is presented in [18]. In the work, the novel approach produced an excellent result superior to RBF, PNN and MLP networks. [19] presented a hybrid genetic-neural (GA-ANN) model to differentiate malignant from benign in a group of patients with histopathologically proved breast lesions in the base of BI-RADS descriptors and data derived from time-intensity curve. The study shows that a good accuracy of 91%, sensitivity of 95% and specificity of 78% is yielded compared to the radiologist's opinion. [20] presented an overview of the current research being carried out using the data mining techniques to enhance the breast cancer diagnosis and prognosis. From the work, it is observed that the accuracy for the diagnosis analysis of various applied data mining classification techniques is highly acceptable for professional decision making for early diagnosis and avoid biopsy. However, more efficient models can provided for prognosis problem by inheriting the best features for the models defined. In the study, the best model can be obtained after building several different types of model, or by trying different technologies and algorithms. In [21] a GA-based feature selection method in conjunction with neural network model and statistical classifier were investigated to classify micro-calcification patterns in digital mammograms. In the study, the obtained results show that the proposed approach is able to find an appropriate feature subset and neural classifier achieves better results than the two statistical models. [22] presents a study on classification of breast cancer using feed forward artificial neural networks. The performance of the network shows high accuracy rate of 99.28% using Levenberg-Marquardt training algorithm. In [23], the use of deep max-pooling convolutional neural networks to detect mitosis in breast histology images. The approach outperforms other approaches by a significant margin and won the ICPT 2012 mitosis detection competition. In [24], a novel approach for cancer detection in MRI mammogram using decision tree induction and BPN is presented. In the study, the accuracy of the genetic algorithm was significantly higher than the average predicted accuracy of 0.9612. In [25] the diagnosis of breast cancer using a combination of genetic algorithm and artificial neural network in medical infrared thermal imaging was presented. The result in the study indicates the improvement in the capacity and power of globalizing the ANN in obtaining more accurate and good precision in cancer diagnosis. [4], and [26] carried out survey on the use of various neural network techniques for the classification of breast cancer data, and concluded that the use of ANN increases the accuracy of most of the methods and reduces the need for the human experts. According to [27], the main reason for the use of these different techniques is to guide researchers to develop most cost effective and user friendly systems, processes and approaches for clinicians. In the study,

they opined that the accuracy of neural network can be further enhanced by increasing the number of neurons in the hidden layer and applying different training and learning rules for training ANN in order to improve the performance of the classifiers.

2. Methodology

The data used in the present work was obtained from the University of California at Irvine (UCI) Machine Learning Data Repository. The dataset was downloaded from their website. The downloaded file contains medical data concerning breast cancer classification cases that were categorized by medical experts to malignant or benign. The downloaded dataset contains features that describe characteristics of the cell nuclei of a Fine Needle Aspirate (FNA) of a breast mass.

3.1 Description of Data set: Number of instances 699; Number of attributes: 10 plus the class attribute; Attributes 2 through 10 will be used to represent instances; Each instance has one of 2 possible classes: benign or malignant; Class distribution: Benign: 458 (65.5%), Malignant: 241 (34.5%)

3.2 Attribute Information:

Attribute	Domain
1. Sample code number	id number
2. Clump thickness	1-10
3. Uniformity of cell size	1-10
4. Uniformity of cell shape	1-10
5. Marginal adhesion	1-10
6. Single epithelial cell size	1-10
7. Bare nuclei	1-10
8. Bland chromatin	1-10
9. Normal nucleoli	1-10
10. Mitosis	1-10
11. Class	(2 for benign, 4 for malignant)

The dataset was then divided into training, testing and validate sets in the ratio of 70:15:15 respectively. The training set is to fit the model i.e. it is used for computing the gradient and updating the network weights and biases. The validation set, known as a pseudo test set, is to evaluate the quality of the model during training. It is used for monitoring the error during training, i.e., during training, the error on the validation data set decreases until over-fitting starts to occur at which point the error starts to increase again. While the test set, out-of-sample test sets, measures the performance of the resulting network, i.e., testing how accurate the network is. The dataset is scaled between the upper and lower bounds using the Max-Min function in Equation 1. The scaling was carried out so as to map the desired range of variables ranging between the minimum and maximum range for the network use. The dataset contains 10

attributes (1 class and 9 numeric features). The 9 numeric features are in the analog form scaled in the range between 0 and 1 using Equation 1 for machine learning.

$$X' = (max_2 - min_2) \frac{x - min_1}{max_1 - min_1} + min_2 \quad \dots (1)$$

$$number\ of\ hidden\ units = 2n + 1 \quad \dots (2)$$

where n is the number of input, X' is the normalized value, x the original value, min₁ and max₁ are the minimum and maximum values of all original values respectively, and min₂ and max₂ are the expected minimum and maximum of the new scaled values.

To accomplish the task of training the networks, the training dataset was first executed using the three neural networks- one multilayer perceptron neural network and two deep neural networks using 70% of the dataset. The number of hidden neurons was calculated using Equation 2 as suggested by [28] and [29] as one possible way of determining the hidden units. However, to extend the size of the 2 and 3 hidden layers, the number of the first hidden units obtained is divided by 2 to determine the second hidden units. Then for the 3 hidden layers, the number of the first hidden unit is divided by 4. That means half of the size of the first hidden units determine the size of the second layer and half of the second hidden layer determines the size of the third hidden units respectively. Thereafter, the models obtained are then validated using 15% of the validation dataset. The validated model was then tested using the testing dataset to know how well the model can generalize on unknown instances. The normalized data set fed into the networks represent 0-1 binary values. The target matrix included 2 classes: benign and malignant cases. In the case where the cancer type matched the class of the column, the value corresponding to the value in the row would be 1 and the other row would be 0. The training of the network process was repeated for up to 500 epochs. This work was programmed in Python programming language using Jupyter notebook on Anaconda Navigator IDLE software. The libraries imported for the implementation are pandas, numpy, matplotlib, sklearn, torchvision, torch, seaborn.

3. Results and Discussion

The performance of the different models and the existing based model were compared and are shown in Table 1. The details are given in respect to their sample sizes and classification efficiencies i.e. the number of layers. The proposed models all showed higher accuracies of about 95.23% compared to the existing model with 92% using a multilayer and 88.9% when a single layer was used respectively. The existing model used a training sample of 350 datasets while the proposed model used 699 samples respectively. The proposed model used one, two and three hidden layers and achieved this feat because of the way the model is designed and the choice of parameters used; while the existing model used one and two respectively. To investigate the degree of successfulness and applicability of the models, the study used the confusion matrix, the Area under ROC curve and other performance metrics. The confusion

matrix is a table layout showing the performance of a supervised learning algorithm in a visualized form and it is the preferred performance measure for classifier system. Some of the matrices adopted are:

- a.) Sensitivity (True Positive Rate or Recall): means positive classes correctly classified as positive classes and represented as:

$$\frac{TP}{(TP+FN)} \quad \dots (3)$$

- b.) Specificity (True Negative Rate): measures the negative instances truly labeled as negative by classifier. Specificity should be high and given as:

$$\frac{TN}{(TN+FN)} \quad \dots (4)$$

- c.) Precision: shows the ratio of the total number of correctly classified positive examples and the total number of predictions that are correct, and given as:

$$\frac{TP}{(TP+FP)} \quad \dots (5)$$

- d.) Accuracy: measures the proportion of the total number of predictions that are correct and given as:

$$\frac{TP+TN}{(TP+TN+FP+FN)} \quad \dots (6)$$

- e.) F-Measure (or F-score): measures the weighted average of the true positive rate (recall) and precision. The F-score reaches its best value at 1 (perfect precision and recall), and worst at 0. It is a measure of a test's accuracy and given as

$$2 * \frac{Precision * Recall}{Precision + Recall} \quad \dots (7)$$

Where the TP, FP, FN and TN stands for the True positive, False positive, False negative and True negative respectively.

The receiver operating characteristic (ROC) curve is another common tool used to evaluate binary classifiers [30]. ROC graph is a technique for visualizing and selecting classifiers based on their performances. ROC graph summarizes the performance of a classifier over all possible thresholds (a numeric value that represents the degree of which an instance is a member of a class). The performance measure of the first model with single hidden layer and details of the results obtained are shown in Table 1.

Table 1: Performance analysis of variant classifiers

Measures	Proposed models			Existing model	
	1 st Model	2 nd Model	3 rd Model	Single	Multilayer
Accuracy	0.9523	0.9523	0.9523	0.889	0.92
Sensitivity	0.9565	0.9565	0.9565	-	-
Specificity	0.0434	0.0434	0.0434	-	-
Precision - Train	1.00	1.00	1.00	-	-

-Test	0.93	0.95	0.95	-	-
F-Score				-	-
-Train	1.00	1.00	1.00		
-Test	1.00	1.00	1.00	-	
Sample size	699	699	699	350	350

The various performances show the behaviour of both the training and testing datasets graphically with respects to convergence rate, MSE loss functions, and the AUROC curve respectively. The MSE loss functions of the variant models are shown in Figures 1-3. The convergence rates of the three models are depicted in Figures 4-6. The ROC curve showing the true positive rate and false negative rate for the three models are depicted in Figure 7-9 respectively while the curve for the confusion matrix is depicted in Figure 10.

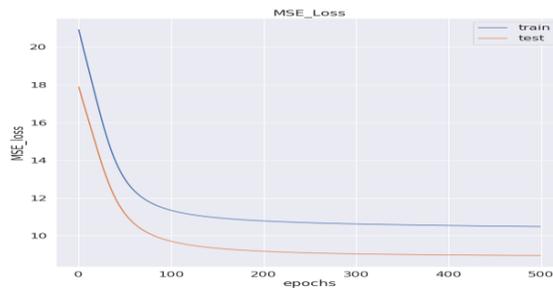


Figure 1: MSE Loss function for single hidden layer MLP

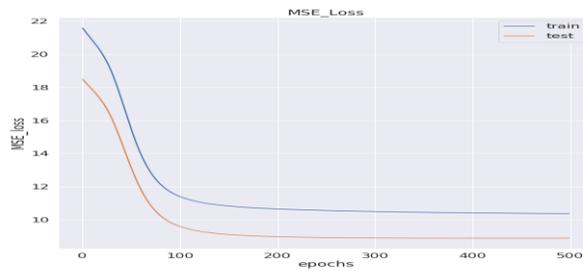


Figure 2: MSE Loss function for two-hidden-layer deep NN

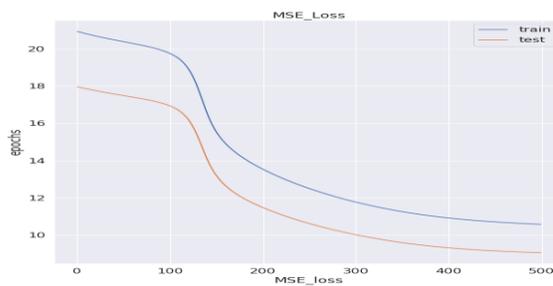


Figure 3: MSE Loss function for three-hidden-layer deep NN

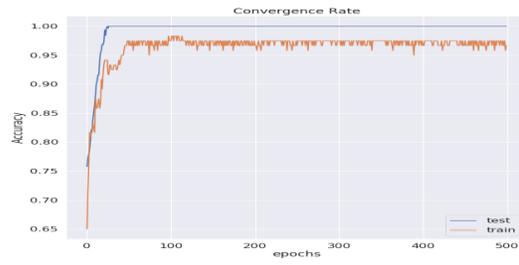


Figure 4: Convergence rate for the 1st model

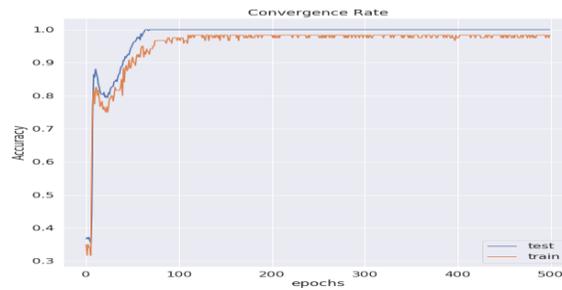


Figure 5: Convergence rate for the 2nd model

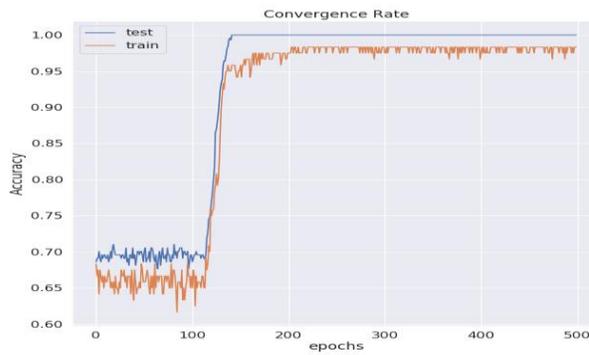


Figure 6: Convergence rate for the 3rd model

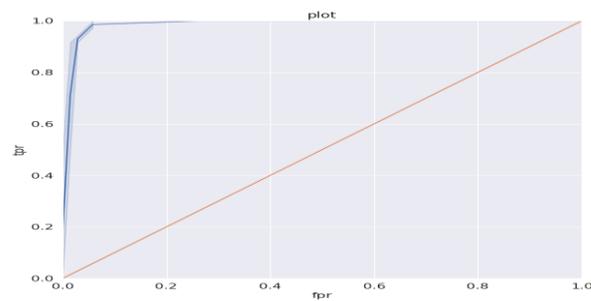


Figure 7: ROC curve for 1st model

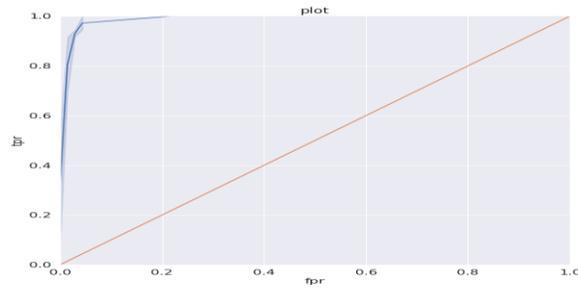


Figure 8: ROC Curve for 2nd model

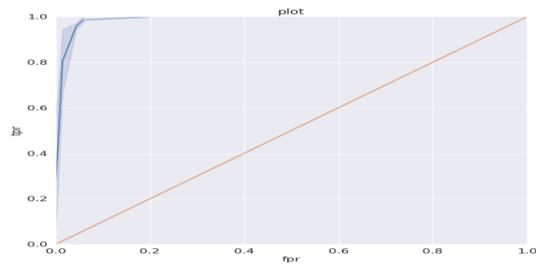


Figure 9: ROC curve for the 3rd model

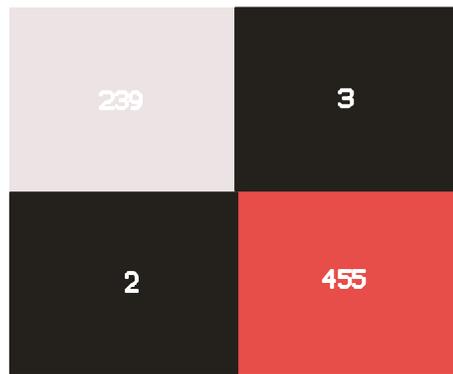


Figure 10: Confusion matrix

The distinct behaviours of the three models can be seen vividly from the convergence rate showing in Figures 4, 5 and 6 respectively. It simply explains the fact that all three models converges at different rates. While it took less time for the first single layer model, and slight different from the second model, it took quite some amount of time for the third model (i.e. the deep NN). This is due to the size of the architecture. Also, the slope on the MSE loss for the first model in Figure 1 is slightly distinct from the second and third model in Figures 2 and 3 respectively. The MSE loss of the third model shows a sigmoid shape indicating an unsteady slope to the first and second models. Figure 10 shows that the only five instances were misclassified which is an improvement of the model developed.

4. Conclusion

In this study, three variant artificial neural networks was used to classify the types of breast cancer i.e. benign and malignant. The networks were trained first, then validated and tested using the various datasets. The various parameters of accuracy, sensitivity, and specificity values were obtained which appears to be all the same except for the test precision value found to be 0.93, 0.95 and 0.95 from the first, second and third models respectively. The resultant values show a high acceptability level of reliability in the cases of classification of the dataset. Models also demonstrate the use of appropriate training parameters to achieving high performance irrespective of the depth of the network architecture adopted. It is evident that one of the reasons for this is the normalization of the input vectors and the choice of training parameters adopted. In fact, the models implemented in this work can be better improved in the future if other model like evolutionary algorithms are combine in the selection of optimal parameters for better performance. We strongly believe if these models are hybridized using evolutionary algorithm the epochs will be reduced producing high accuracies and better performance.

References

- [1] Paulin, F. and Santhakumaran, A. (2010) Back propagation neural network by comparing hidden neurons: case study on breast cancer diagnosis. *Int. Journal of Computer Application*, Vol. 2(4), pp. 40-44.
- [2] Adeli, M. and Zarabadipour, H. (2011) Automatic disease diagnosis system using pattern recognition based genetic algorithm and neural networks. *Recent Advances in Applied & Biomedical Information and Computational Engineering in Systems Applications*, pp.36-41.
- [3] Huang, Y. Li, W and Ye, X. (2011) A study of Genetic neural network as classifiers and its application in breast cancer diagnosis. *Journal of Computers*, Vol.6(7), 1319-1324.
- [4] Saxena, S. and Burse, K. (2012) A survey on neural network techniques for classification of breast cancer data. *International Journal of Engineering and Advanced Technology (IJEAT)*, Volume 2, Issue 2, pp. 234-237.
- [5] Dubey, A.K. and Jain, U.G. (2015) Breast cancer statistics and prediction methodology: A systematic review and analysis, *Asian Pac J Cancer Prev*. 2015; 16(10):4237-45
- [6] Thein, H.T.T., and Tun, K, M.M. (2015) An approach for breast cancer diagnosis classification using neural network. *Advanced computing: An international Journal (ACIJ)*, Vol. 6, No. 1, pp. 1-11
- [7] Rani, K.U. (2010) Parallel approach for diagnosis of breast cancer using neural network technique. *International Journal of Computer Application*, Vol. 10(3), pp. 1-5.
- [8]. Saini, S.S and Vijay, R. (2014) Performance analysis of artificial neural network based breast cancer detection system. *International Journal of Soft-computing and Engineering (IJSCE)*, Volume 4, Issue 4, pp. 70-72.
- [9]. Venkatesan, P. and Premalatha, V. (2012) Genetic-neuro approach for disease classification. *International Journal of Science and Technology*, Volume 2, No. 7, pp. 473-478.
- [10] Mohamed, M.A., Hegazy, A.E-F., Badr, A.A. (2011) Evolutionary fuzzy ARTMAP approach for breast cancer diagnosis, *IJCSNS International Journal of Computer Science and Neural Network Security*, Vol., 11, No. 4, pp77-84.
- [11] Kalteh, A.A., Zarbkhsh, P., Jirabadi, M., and Addeh, J. (2015) A research about breast cancer detection susign different neural networks and K-MICA algorithm. Accessed on 3rd August, 2015 from <http://www.cancerjournal.net>

- [12] Ahmed, A.M., Muhammad, G., and Miller, J.F. (2012) Breast cancer detection using Cartesian genetic programming evolved artificial neural networks. Accessed from https://www.researchgate.net/publication/228548271_Breast_Cancer_Diagnostic_Factors_Elimination_via_Evolutionary_Neural_Network_Pruning/citations
- [13] Balanica, V., Dumitrache, I., and Preziosi, L. (2013) Breast cancer diagnosis based on speculation feature and neural network techniques, *Int., J. Comput., Commun.*, 8(3): 354-365
- [14] Adamopoulos, A.V. (2007) Breast cancer diagnosis factors elimination via evolutionary neural network pruning
- [15] Bevilacqua, V., Mastronardi, G., Menolascina, F., and Pedone, P.A. (2006) A Novel multi-objective genetic algorithm approach to artificial neural network topology optimization: the breast cancer classification problem. *International Joint Conference on Neural Networks Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada*, 3916-3923
- [16] Khosravianian, A., and Ayat, S. (2016) Diagnosis of breast cancer type by using probabilistic neural network in decision support system. *International Journal of Knowledge Engineering*, Vol., 2, No. 1, pp. 73-76.
- [17] Adam, A., and Omar, K. (2018) Computerized breast cancer diagnosis with genetic algorithms and neural network, - fitt.mmu.edu.my/caic/papers/afzaniCAIET.pdf.
- [18] Punitha, A., Sumathi, C.P., and Santhanam, T. (2007) A combination of genetic algorithm and ART neural network for breast cancer diagnosis, *Asian Journal of Information Technology* 6(1): 112-117.
- [19] Nirooei, M., Abdolmaleki, P., Tavakoli, A., and Gity, M. (2009) Feature selection and classification of breast cancer on dynamic magnetic resonance imaging using genetic algorithm and artificial neural network, *J. Electrical System*, 5-1.
- [20]. Gupta, S., Kumar, D., and Sharma, A. (2011) Data mining classification techniques applied for breast cancer diagnosis and prognosis, *Indian Journal of Computer Science and Engineering (IJCSE)*, Vol.2(2), pp. 188- 195
- [21]. Zhang, P., Verma, B., and Kumar, K. (2004) Neural vs. statistical classifier in conjunction with genetic algorithm based feature selection, *Pattern Recognition Letters*, Elsevier
- [22]. Paulin, F., and Santhakumaran, A. (2011) Classification of breast cancer by comparing back propagation training algorithms, *International Journal on Computer Science and Engineering (IJCSE)*, 3(1), pp. 327-332
- [23]. Cirezan, D.C., Giusti, A.G., and Schmidhuber, J.G. (2013) Mitosis detection in breast cancer histology images with deep neural networks, In *Proc. MICCAI*, volume 2, pp 411–418.
- [24]. Saravanan, V., and Angayarkanni, S.P. (2010.) A novel approach for cancer detection in MRI mammogram using decision tree induction and BPN, *International Journal of Image Processing (IJIP)*, Volume 4, Issue 6, pp. 661-668
- [25]. Zadeh, H.G., Haddadnia, J., Hashemian, M., and Hassanpour, K. (2012) Diagnosis of breast cancer using a genetic algorithm and artificial neural network in medical infrared thermal imaging. *Iranian Journal of Medical Physics*, Vol., 9, No, 4 pp. 265-274.
- [26] Beg, M.M., and Jain, M. (2012) an analysis of the method employed for breast cancer diagnosis, *International Journal of Research in Computer Science*, vol. 2, no. 3, pp. 25-29
- [27] Agrawal, S., and Agrawal, J. (2015) Neural network techniques for cancer prediction: A survey, *Procedia Computer Science* 60, 769-774, Elsevier
- [28] Sivanandam, S.N, Sumathi, S., and Deepa, S.N. (2014) *Introduction to Neural Networks Using Mat lab* Springer
- [29] Kokko, T. (2013) *Neural Networks for Computationally Expensive Problems*. Master Thesis.
- [30] Geron, A. (2019). *Hands on Machine Learning with Scikit-Learn, Keras and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 2nd Edition, O'Reilly.