

Prediction of Breast Cancer Type Based on Artificial Intelligence Technique

Shameem Hasan

M.Sc. Engineering student

Department of Electrical and Electronic Engineering

Islamic University of Technology

Gazipur-1704

Bangladesh

ABSTRACT

The fundamental cause of death among women in developing nations of the world is breast cancer. It is very common and is considered the second most dangerous disease all over the world due to its death rate. Recently, there has been a dramatic increase in breast cancer cases among women. It is known to us that at the time of mutation, cells get an ability to go on dividing without control and results in forming a tumor. This tumor can be of benign or malignant in nature. The benign tumors are not dangerous but malignant tumors are dangerous to health. The unchecked malignant tumors have the ability to spread in other parts of the body. At the primary stage, breast screening is a method of detecting breast cancer. The first step involves taking an x-ray of each breast which is called a mammogram. The mammogram can detect small changes in breast tissue which may indicate cancers which are too small to be felt. Artificial Neural Network is an effective tool that has found application in the field of medical diagnosis for early detection and diagnosis of breast cancer. If breast cancer is detected in early stage, chances of survival are comparatively high.

Key Words: Artificial Neural Network, Malignant tumor, Benign tumor, Mammogram, Mean Square Error.

1. INTRODUCTION

Cancer is the type of disease in which cells get changed and keep growing out of control. Most types of cancer cells form a lump or mass called tumor and are named after the part of the body where the tumor first starts [1]. The tumor cannot be referenced as cancer. Micro calcifications and masses are the earliest signs of breast cancer. Micro calcifications are clusters of calcium deposits which are very small in size and present inside the soft breast tissues [2]. Generally, detection of lumps or masses in breast tissues is more challenging compared to the detection of micro calcifications. Masses often exhibit poor image contrast when using mammography [3].

Breast cancer occurs in both men and women, but mostly in women. Breast cancer is the most common cancer among women excluding skin cancers. This cancer affects one in eight women during their lives. Although scientists know some of the risk factors (i.e. aging, genetic risk factors, family history, menstrual periods, obesity etc.) that increase a woman's chance of developing breast cancer, but they do not yet know what causes most breast cancers or exactly how some of these risk factors cause cells to become cancerous.

The death rate due to breast cancer is very high over the world. According to WHO (World Health Organization) breast cancer has an impact on over 1.5 million women each year around the world. In 2015, 570,000 women died due to breast cancer which is approximately 15% of all death among women from cancer [4]. The death rate due to breast cancer can be reduced by following proper diagnosis technique at an initial stage. Various techniques have been used for the detection of breast cancer. Mammography is very effective and most commonly used technique for the early detection of breast cancer [6-10].

The difficulty in the classification of benign and malignant type breast cancer causes a significant problem. Consequently, automated classifiers may be useful for radiologists in distinguishing between benign and malignant patterns. In this paper, Artificial Neural Network (ANN) which can be served as an automated classifier for the prediction of the type of breast cancer is proposed.

2. ARTIFICIAL INTELLIGENCE (AI) IN BREAST CANCER PREDICTION

Artificial intelligence (AI) is a rapidly growing branch of technology. The AI, with the help of computer science and statistics, allows computers to learn without explicit programming through automatic extraction and analysis of complex data. Artificial Intelligence algorithm is a medium to discover new materials, master various games and improve predictive ability in clinical medicine. This AI strategy can be applied to enhance the accuracy of prediction associated with cancers that were missed in routine screening practice, further enhancing the feasibility and potential role of AI in health research [11-16].

3. STATISTICAL ANALYSIS

In this research work, many statistical approaches such as Bayesian regularization (BR), Levenberg–Marquardt algorithm (LM), and scaled conjugate gradient (SCG) have been used for the evaluation and accuracy of the performance and results. The data are taken from internet sources. Bayesian classification is the technique to construct the classifiers. Classifiers are nothing but the models that assign the class labels to the problem instance. Levenberg–Marquardt algorithm is also known as DLS that is damped least squares used for solving generic curve fitting problems by finding the local minimum [17]. Scaled conjugate gradient is feed-forward and supervised algorithm for neural networks. Feedforward here means that in connections there is no loop between the units. The general equations corresponding to each are mentioned below:

Bayesian regularization:

$$x = \arg_{b \in \{1, \dots, B\}} \max p(C_b) \prod_{i=1}^n p(y_i | C_b) \tag{1}$$

Levenberg–Marquardt algorithm:

$$H(\beta) = \sum_{j=1}^m [x_j - f(y_j, \beta)]^2 \tag{2}$$

Scaled conjugate:

$$S_k = \frac{\dot{E}(W_k + \sigma_k P_k) - \dot{E}(W_k)}{2} \tag{3}$$

Bayesian regularization, Levenberg–Marquardt algorithm and scaled conjugate are the various algorithmic parameters and functions used in the neural networks. BR can eliminate or reduce the need for lengthy cross-validations and it is more robust than the standard back-propagation methods, whereas to solve nonlinear least squares problems, the LM technique is considered to be the standard one as it shows a lower performance in terms of predictive ability. On the other hand, SCG needs O(n) of memory where n represents the number of weights in the network although it uses second order of information from neural networks [17,18]. Among these three, the BR is considered to be the optimal one as it develops the nonlinear relationships and it has more predictive abilities. To get better and refined results, the data was tested through twenty hidden layers, and on observing the results, it can be seen that the BR shows least mean square error (Table 1) for breast cancer prediction.

Table 1. No. of Hidden layers, Mean Square Error (MSE) and No. of Iterations for breast cancer prediction

Serial No.	No. of hidden layers	Mean square error	No. of iterations
1	10	0.022318	413
2	15	0.024745	412
3	20	0.020012	340
4	25	0.02873	529
5	30	0.027209	96

3.1 NEURAL NETWORKS

It is one of the concepts that has been inspired by the functionality of the human brain and its performance in the identification of phenomena. Neurons (a single neuron is shown in Figure 1) are placed in different layers in multilayer neuron network. Input layer being the first layer receiving information and till its capability with other neurons, it transfers the information in the form of input signals to the other next layers. Neuron weight is the communication ability of each neuron with other neurons. The number of neurons in each layer depends on the weight of neuron and the previous layers' neurons. In addition to the input layer, the neural network also consists of the hidden layers and the output layers. Some of the advantages of using artificial neural networks(ANN) are its immaculate on-point accuracy on a wide variety range of problems, less requirement of formal statistical training, offering various multiple training algorithms and having the implicit ability to detect nonlinear complex relationships between independent and dependent variables. In this, a neuron is the main processor and adding neurons to hidden layers will reduce calculation error but will be more time to consume for calculations [20].

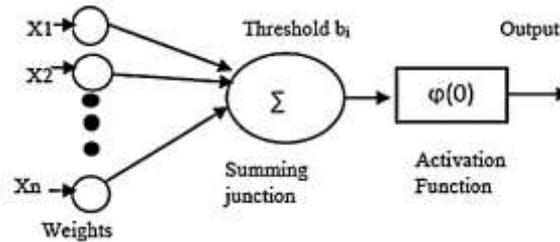


Figure 1: Single Neuron

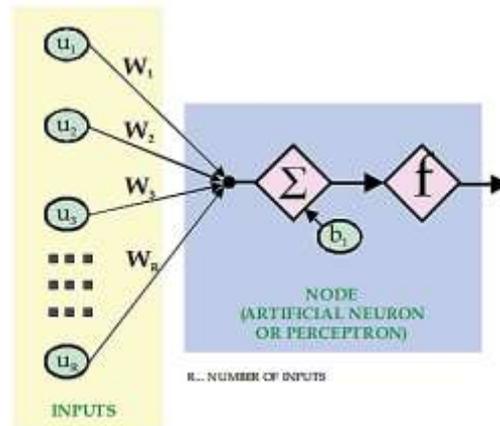


Figure 2: Node (Artificial neuron or perceptron)

The neural network has 10 input layer, several hidden and one output layer. The input parameters are the values of different effective factors of patients, the auto regression component and the physical variables for which the statistical analysis has shown the strongest influence (Figure 3). The number of auto regression components m is determined by the correlation coefficients between the predicted value and its previous values persistence of the predicted variable. The goal is forecasting one step ahead. The number of neurons in the first and second layers is determined by the criterion of the minimum squared error(MSE).

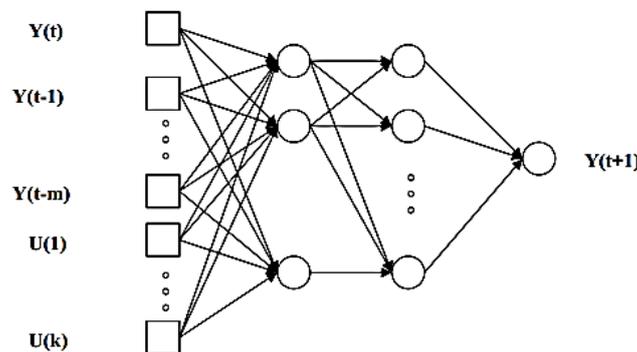


Figure 3: The architecture of the Neural Network

4. DATA AND METHODOLOGY

The datasets include 10 input parameters. The parameters are:

- Mean Radius – Mean of distances from the center to points on the perimeter.
- Mean Texture – Standard deviation of grayscale values.
- Mean perimeter – Mean size of a core tumor.
- Mean area – Mean area of a core tumor.
- Mean smoothness – Mean of local variation in radius lengths.
- Mean compactness – $(\text{Mean of perimeter})^2 / \text{area} - 1$
- Mean concavity – Mean of the severity of concave portions of the contour.
- Mean concave points – Mean for no. of concave portions of the contour
- Mean symmetry
- Mean fractal dimension – Mean for “coastline approximation” - 1

The methodology of the whole procedure is shown in Figure 4 with the help of a flowchart for a better understanding.

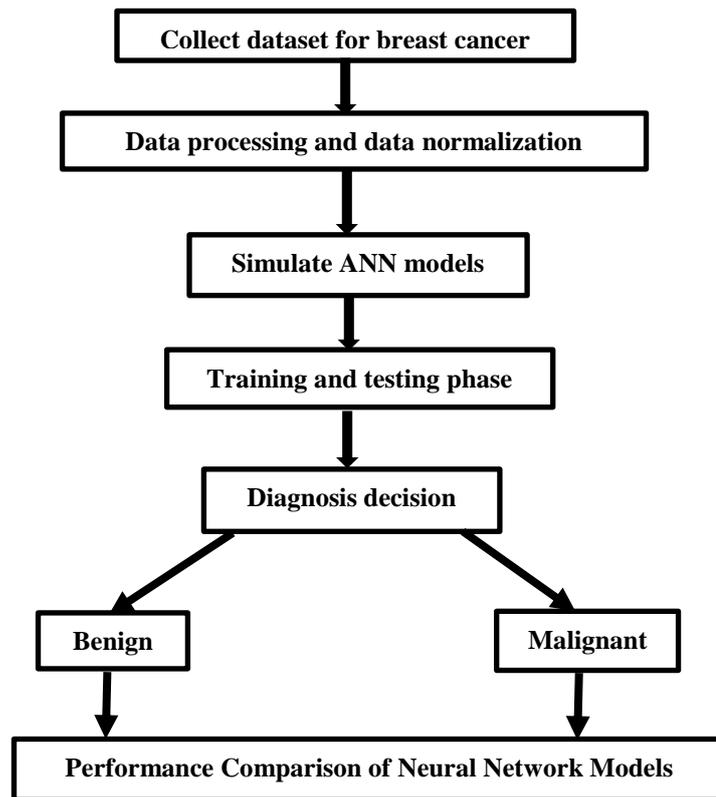


Figure 4: Methodology of work

There are three portions of the data that act as a training set, validation set and testing set in neural networks having 70%, 15%, and 15% weight respectively. For this paper, 569 sample data of different patients have been used collected from internet sources. Each record in the database has ten attributes. In this database, 215 (37.79%) records are malignant and 354 (62.21%) records are benign.

The nftool (neural fitting tool) of MATLAB has been utilized in our work to determine the performance and the results. In this tool, the number of input data to layer and the number of hidden layers have to be defined. In this paper 10, 15, 20, 25 and 30 hidden layers have been used one by one to get the output and their mean square errors have been compared to get the performance measures. Since layer 20 gives the most refined results with the least value of the error, the data has been trained till 20 layers.

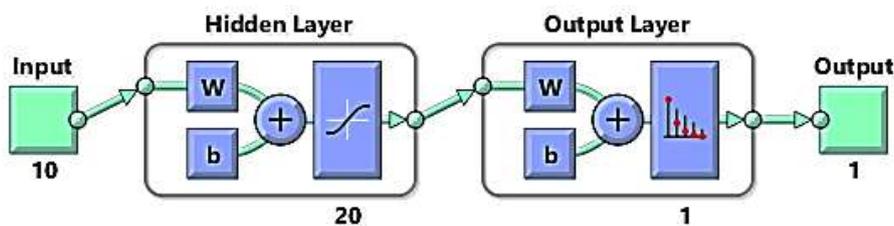


Figure 5: Optimum Neural Network working for breast cancer prediction

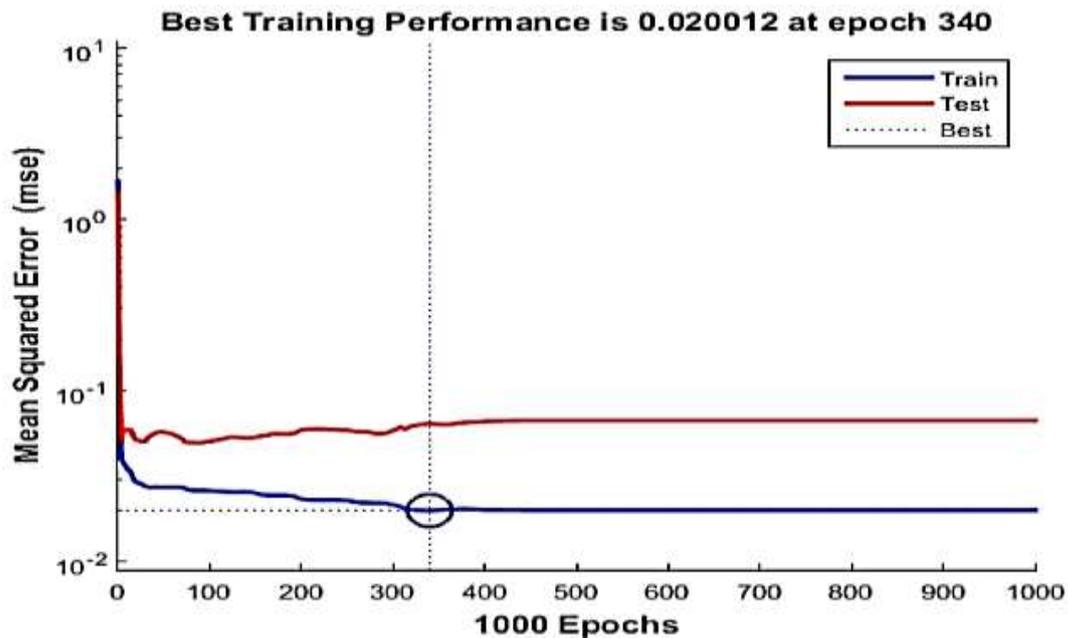


Figure 6: Performance graph for breast cancer type prediction showing mean square error

The neural network is given with the inputs such as the values of number of times pregnant, glucose level, blood pressure measurement, skin thickness, amount of insulin, BMI, diabetes pedigree function, age, the hidden layers, and the output having three desired values such as for malignant type breast cancer (0) and benign type breast cancer (1). These are levels of the prediction in which the range of its values 500 and above is considered to be serious, range from 250 to 500 is medium, and 0–250 is safe. Table 1 shows the number of hidden layers with their mean square.

Figure 6 shows that the validation, test data sets and the performance of training with respect to epochs of breast cancer prediction. As it is clearly seen that the best performance measure is at Serial No. 3 i.e., with the 20 hidden layers and with minimum mean square value i.e. 0.020012 for prediction. Figure 7 shows the 2D plotting of the input and desired output variables for prediction.

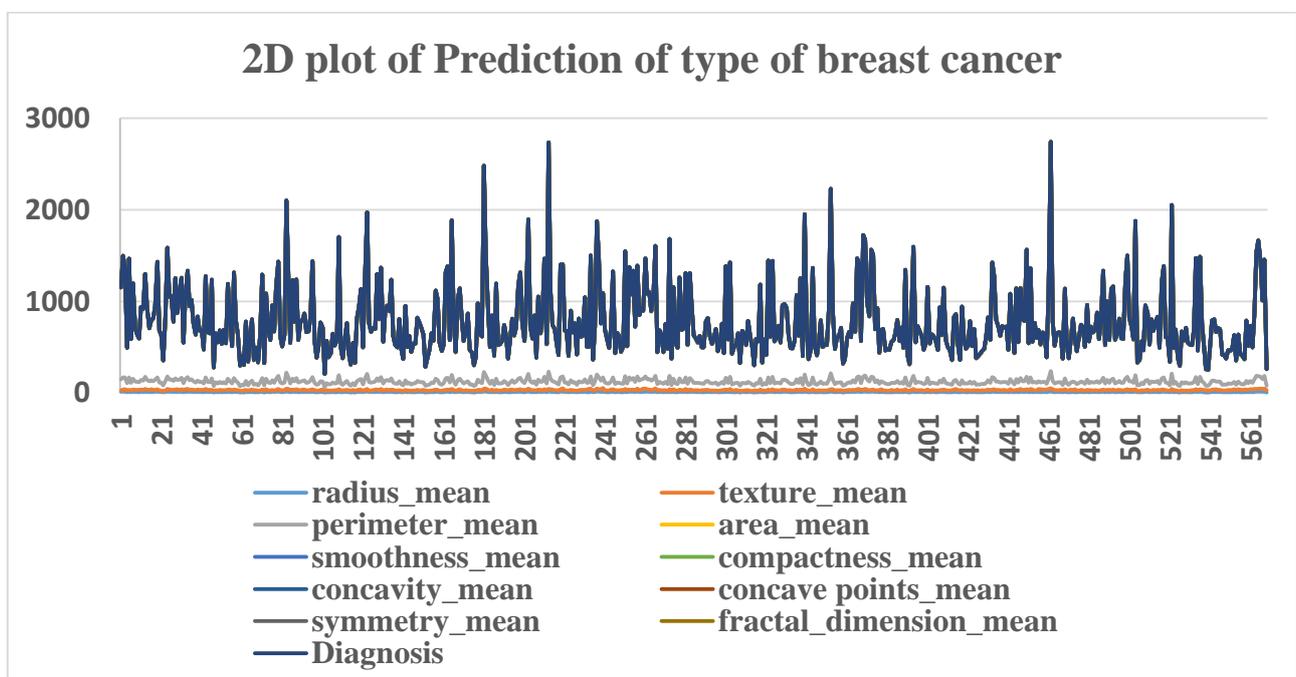


Figure 7: 2D plot of prediction of breast cancer type.

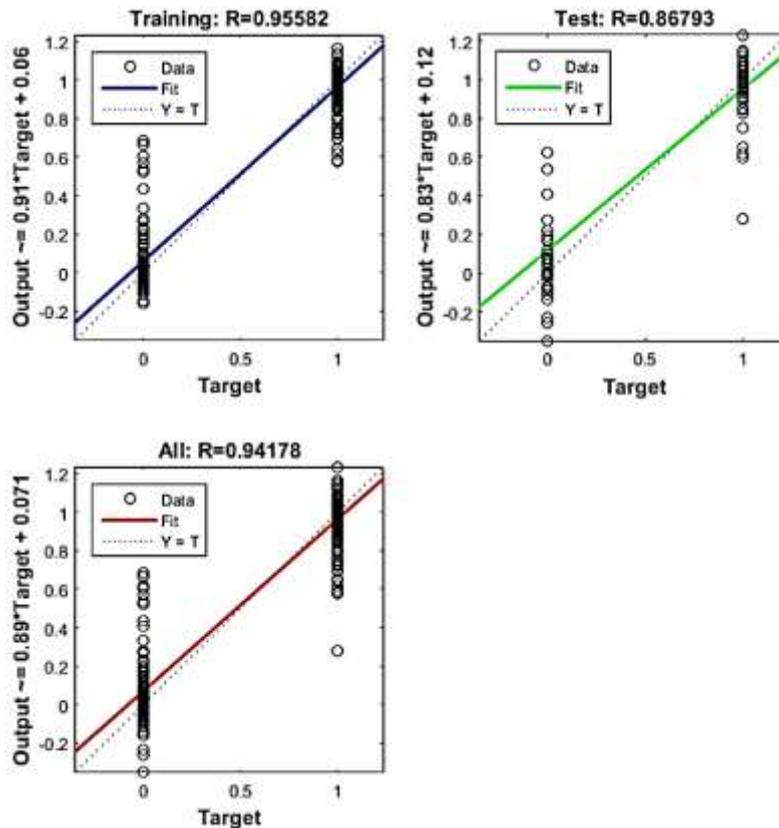


Figure 8: Regression curve for breast cancer type prediction

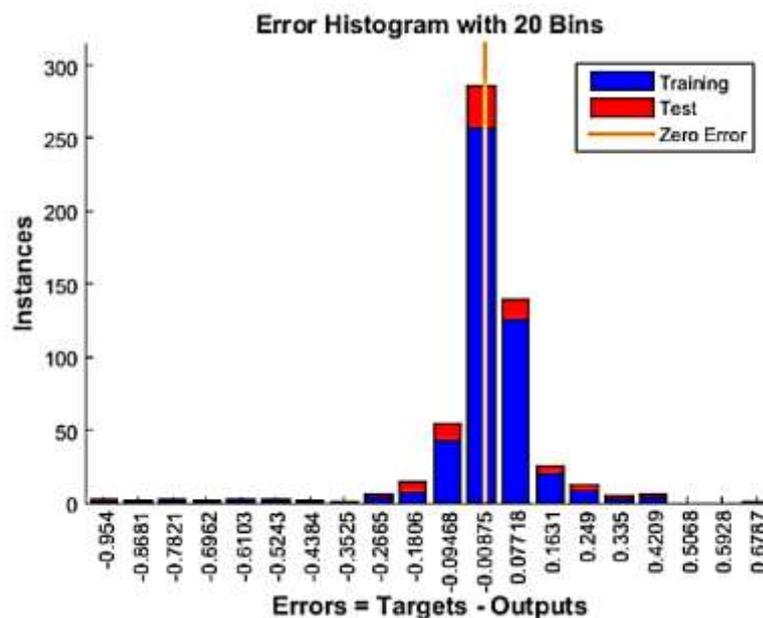


Figure 9: Error histogram for breast cancer type prediction

Further, Figure 8 and Figure 9 show the regression curve between the target data and the output of levels of the breast cancer prediction and the error histogram of the analysis respectively.

5. REGRESSION ANALYSIS

Regression analysis is a statistical process to estimate association among all variables. In the regression plot output from the network are plotted versus the target set shown in Figure 8. In the regression plot, a perfect fit is indicated by dotted line while the solid line shows the output from the network. Solid line perfectly equal to a dashed line is achieved if the classifier predicts 100% accurately. The difference between two line shows there is some sample which is not correctly predicted by the network. Data is

represented by a circle. In the plot shown below the value of R is 0.94178, this value also shows the result accuracy. The confusion matrix is shown in Figure 10 for specificity, sensitivity, and accuracy of different data division.

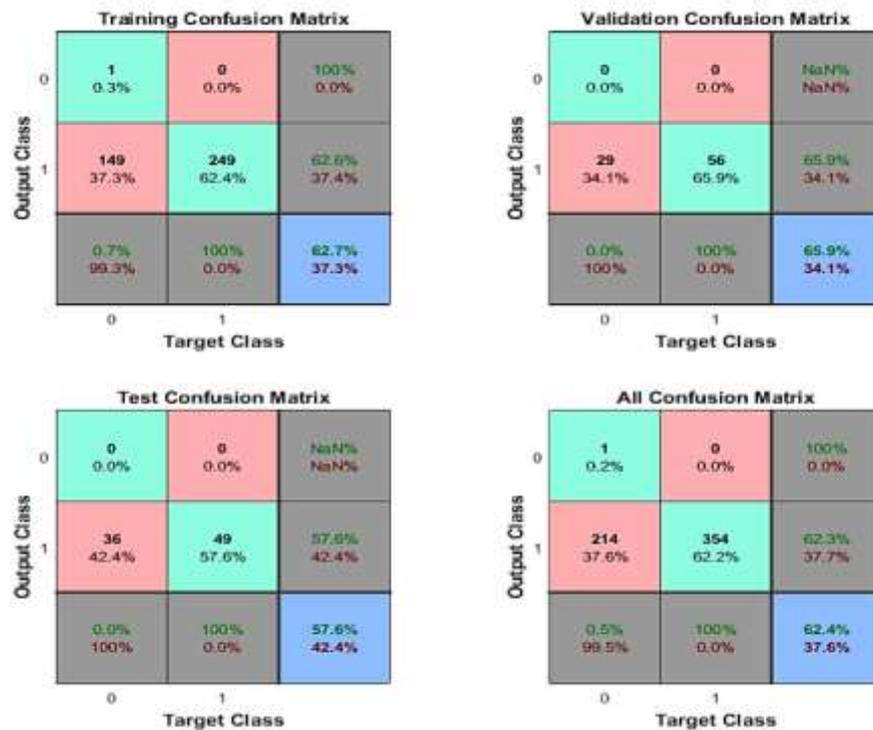


Figure 10: Confusion matrix for training, test and validation data.

6. CONCLUSION

In this paper, the predictive analysis of breast cancer has been done through artificial neural networks. The different parameters of the neural network have been used to get accurate results. Breast cancer has become a major health care problem for women over the world. Interpretation of mammography for breast cancer screening can provide benefit through early detection but can also miss cancer that is present or fast-growing or can result in false-positives. On the other hand, Artificial intelligence (AI) is capable of advanced learning using large complex datasets. In fine, it can be concluded that ANN is effective in prediction of breast cancer type with more precision than other conventional detection methods.

REFERENCES

- [1] L. Hadjiiski, B. Sahiner, M. A. Helvie et al., “Breast masses: computer-aided diagnosis with serial mammograms”, *Radiology*, 2006, vol. 240(2), pp. 343–356.
- [2] T. Balakumaran, I. L. A. Vennila, C. G. Shankar, “Detection of micro calcification in mammograms using wavelet transform and fuzzy shell clustering”, *International Journal of Computer Science and Information Technology*, 2010, vol. 7(1), pp. 121–125.
- [3] J. Dheeba, N. Albert Singh, S. Tamil Selvi, “Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach”, *Journal of Biomedical Informatics*, 2014, 49, pp. 45–52.
- [4] Siegel RL, Miller KD, Jemal A., “Cancer statistics, CA”, *A Cancer Journal for Clinicians. American Cancer Society*, 2016, vol. 67, pp. 7-30.
- [5] Jalalian A, Mashohor S, Mahmud R, Karasf B, Saripan MIB, et al., “Foundation and methodologies in computer-aided diagnosis systems for breast cancer detection”, 2017, vol. 16, pp. 113.
- [6] Singh AK, Gupta B, “A novel approach for breast cancer detection and segmentation in a mammogram”, *Procedia Computer Science*, 2015, vol. 54, pp. 676-682.
- [7] Pereira DC, Ramos RP, Do Nascimento MZ, “Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm”, *Computational Methods Programs Biomedical*, 2014, vol. 114, pp. 88-101.
- [8] Dhahbi S, Barhoumi W, Zagrouba E., “Breast cancer diagnosis in digitized mammograms using curvelet moments”, *Computational Biological Medicine*, 2015, vol. 64, pp. 79-90.
- [9] Muramatsu C, Hara T, Endo T, Fujita H, “Breast mass classification on mammograms using radial local ternary patterns”, *Computational Biological Medicine*, 2016, vol. 72, pp. 43-53.
- [10] Rampun A, Morrow PJ, Scotney BW, Winder J, “Fully automated breast boundary and pectoral muscle segmentation in mammogram”, *Arti Intell. Med.*, 2017.

- [11] Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, et al., “Mastering the game of Go with deep neural networks and tree search”, *Nature*, 2016, vol. 529, pp. 484-489.
- [12] Lu C, Tang X., “Surpassing human-level face verification performance on LFW with Gaussian Face”, *29th AAAI conference on artificial intelligence (AAAI-15)*, 2014.
- [13] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al., “Dermatologist level classification of skin cancer with deep neural networks”, *Nature*, 2017, vol. 542, pp. 115-118.
- [14] Raccuglia P, Elbert KC, Adler PDF, Falk C, Wenny MB, Mollo A, et al., “Machine learning-assisted materials discovery using failed experiments”, *Nature*, 2016, vol. 533(7601), pp. 73-76.
- [15] Ding C, Tao D., “Trunk-branch ensemble convolutional neural networks for video-based face recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.
- [16] Altae-Tran H, Ramsundar B, Pappu AS, Pande V., “Low data drug discovery with one-shot learning”, *ACS Central Science*, 2017, vol. 3(4), pp. 283-293.
- [17] Xingang F, Shuhui L, Fairbank M, Wunsch DC, Alonso E., “Training recurrent neural networks with the Levenberg–Marquardt algorithm for optimal control of a grid-connected converter”, *IEEE Transactions on Neural Network Learn System*, 2015, vol. 26(9), pp. 1900–1912.
- [18] M.I.A. Lourakis, “A brief description of the Levenberg-Marquardt algorithm implemented by levmar”, *Technical Report, Institute of Computer Science, Foundation for Research and Technology - Hellas*, 2005.
- [19] Rasti R, Teshnehlab M, Phung SL, “Breast cancer diagnosis in DCE-MRI using mixture ensemble of convolutional neural networks”, *Pattern Recognition*, 2017, vol. 72 (C), pp. 381-390.
- [20] Abdel-Zaher, Ahmed M, Eldeib, Ayman M., “Breast cancer classification using deep belief networks”, *Elsevier*, 2016, vol. 46, pp. 139-144.