

Breast Cancer Melanoma Prediction using Two Layer Deep Neural Network

Govind Singh
M.Tech Scholar, Dept. of CSE
SIRTS, Bhopal, India

Chetan Gupta
Asst. Prof. , Dept. of CSE
SIRTS, Bhopal, India

ABSTRACT

Breast cancer is one of most commonly diagnosed cancer in the world and second most cause of death after lung cancer. The USA cancer society estimate 284200 women will be diagnosed with BC and 44130 will die due to this disease in 2021. The symptoms of breast cancer include a lump in the breast, bloody discharge from the nipple, chronic pain and changes in the shape or texture of the nipple or breast. Generally BC cancer is classified in two class i. Benign ii. Malign through ML techniques to observe the risk of disease for patient. In this time, Machine Learning (ML) techniques are preferred to classify of BC to achieve best efficiency in diagnoses of BC. In this paper, author wants classification of Breast cancer by using the two layers Deep Neural Network with applying two method gradient & back propagation. Neural Network model generates better result 99.45% as compare to NB classifier 96.19% and KNN classifier 97.51% with minimum error.

Keywords

Breast cancer data, breast cancer classification, deep learning, gradient-descent, epoch, back-propagation, accuracy, ReLU

1. INTRODUCTION

Cancer is one of the most critical disease in the globe today. There are various kind of cancer but breast cancer is one of the most frequent forms among cancer in women. It is second most popular cause of death after the lung cancer. Now a days, there are many advance medical research are explored to find out the disease as early as possible so that it can't spreads in the tissues of the surrounding body [1]. Early identification of cancer is required to enhance the survival rate, & to get benefit of treatment. The latest research has shown that breast cancer survival rate generally from 80 to 90 percent [2] in developing countries while it is much lower in developed countries. Early diagnosis of cruel disease is needed to continue to improve survival with help of Machine learning & medical field research[3] in inclusive way to perform various experiments to identify cancer cells [4]. The ML technique experiment are based on breast's tissue differences by image classification while experiments of medical are based on identifying breast cell differences [5]. There are various classification algorithm of ML to find out malignant breast cancer but most of them having their accuracy less than 99 % [6].

The work analysis from previous findings, suggest improvements to SVMs that are 99% accurate to malignant tumors. The study and results of Support Vector Machine's is useful and valuable in academic because it is most preferred classification algorithms on the Wisconsin BC dataset & also results of this proposal having a lower error rate. Such kind of techniques or algorithm explore new dimension to put on various dataset to find out malignant breast cancer perfectly.

2. LITERATURE SURVEY

There is various type of classification methods are currently present to predict BC using different approach but deep learning is using breast tissue classification method to find critical disease. [8]. The existing study & research on BC are used to predict growth of cancer which is having various features to find impact of disease on patient [9]. Such classification approaches are providing status of patient which is stable or not without previous knowledge. [10]. Machine learning algorithms are tied to present another diagnostic tool with medical test to detect BC .Apart from above, methods of genes expression are also used to identify cancer & in this approach genes are categorized on basis of behavior [11].

The algorithm need to make to find out attributes of patient's laboratory record to identified general systems of malign cancer. [12]. To understand common system of disease, 103564 records have been used to make further direction to other. It is provides feature of BC to detect at early stage, by Zhang et al. (2019) [13]. Such kind of research explores medical understanding of causes, symptoms of breast cancer, and also increases the probability of reliable and prompt diagnoses.

The available classification methods in present time are used analysis of breast cancer in addition to neural networks. Various types of classification method use numeric value to predict BC with respect to existing work. Wisconsin's breast cancer dataset from UCI is most common preferred dataset for classification. In this data set, numerical value between 1 to 10 are available which is related to physical features of BC, & various ML algorithms are applied in different way to achieved higher accuracy for differentiating class of benign and malignant cancer. The malignant cases are to predict by grading task & achieved 97.13 percent accuracy [14]. Using the physical characteristics of BC, multiple classification methods are provide signify result in find out of evil disease cancer.

The accuracy of many of them was evaluated using Wisconsin breast cancer datasets, for example accuracy 94.7 percent was obtained by decision-tree (Quinlan, 1996) & 97.4 percent found by ensemble learning algorithms[15]. For analyzed bosom disease on breast cancer datasets, proposes a deep learning-based scheme with 97.94% accuracy achieved [16]. The supportive vector machines (SVMs) are most widely used method with RBF kernel to make efficient prediction of malignant BC & obtained the 96.8% accuracy [17]. Deep learning assisted Adaboost Algorithm (DLA-EABA) for BC detection has been using CNN-based transfer learning to characterize breast masses for various diagnostic, predictive tasks or prognostic or in several imaging modalities, such as MRI, Ultrasound ,digital breast tomosynthesis and mammography & obtained high accuracy 97.2% [18].

N.Pradhan & V.S. Dhaka [19] have implemented six different machine learning algorithms as DT, NB, LR, RF, SVM, and NN on Wisconsin Breast Cancer dataset and obtained best accuracy of 97.85% through ANN. This paper is suggested that developed model is effectively useful on other data set of BC.

3. PROPOSED METHODOLOGY

The back propagation and gradient descent algorithm are used during implementation of proposed work. The proposed work has been done in three phases. In the first step, author creates a neural network which is trained by using gradient descent algorithm. In the second step, basic network is executed on Wisconsin Cancer Data-set to find out that tumor belonging to particular class of BC using nine different features. At last stage, gradient & back propagation algorithm are applied to find out the local minimum of cost function & also adjusting weight of bias to find out best result with comparing to existing study. The proposed work has been described in figure-1

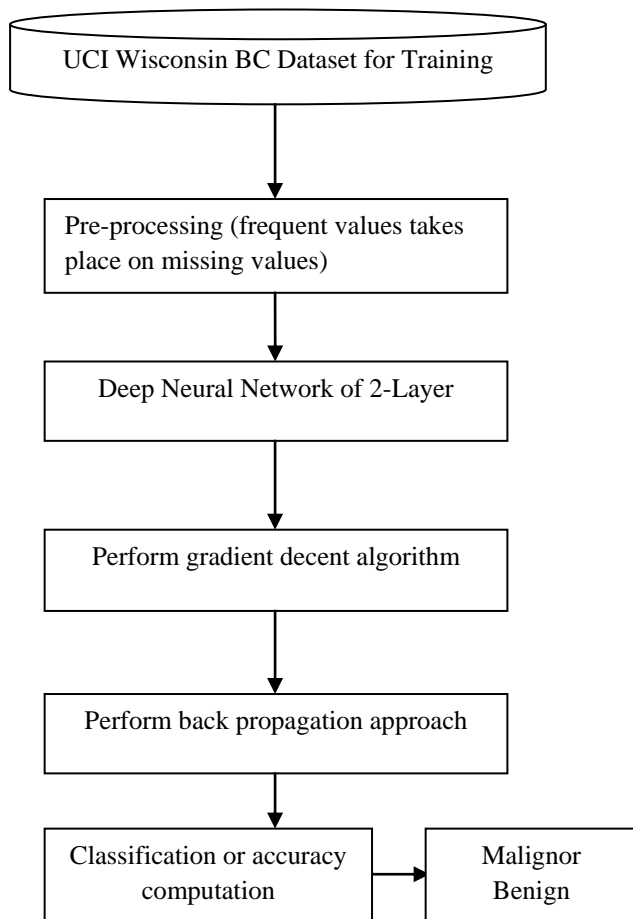


Figure-1 Flowchart of Proposed Work

Dataset-The Wisconsin breast cancer diagnosis dataset is comprises 699 samples with missing values for 16 observations, which limited our dataset to 683 samples initially. This data set having 444 (65%) benign tumors and 239 (35%) malign tumors. It is also mentioned here that data set is consist of 11 attributes in which nine attributes available in Table 1 are used to determine tumor category. ID is left & last feature contains a binary value (2 for benign tumor and 4 for malign tumor).

Table-1 Dataset Feature Details

S. No.	Feature Name
1	Thickness of Clump
2	Cell Size Uniformity
3	Cell Shape Uniformity
4	Adhesion of Margin
5	Cell Size of Single Epithelial
6	Nuclei(Bare)
7	Chromatin of Bland
8	Nucleoli(Normal)
9	Mitoses

Neural Network – In this system, NN has two layers except the input layer.

Input: The deep neural network takes Wisconsin cancer data set as input data source. Proposed network having number of neurons is being equivalent to no. of attributes of our source data.

First layer: it is our first hidden layer which is having a range of secret neurons. The neurons of this hidden layer are bound around them to all units.

Second layer: this is our second hidden layer which is also having similar range of secret neurons as in the first layer. At the end of final out put layer has a one single unit, network output to predict patient suffer from breast cancer or not.

Many researchers suggested that using additional layers to make a 10 or 20 layer network but 2 layer network has been used for convenient to get efficient output. In this model few data have been feed in network's input layer, output of first input layer is used as input to hidden layer then repeat the same for other layer & final output is need to match with input to find desired result.

This system, consist of many unit in network layer which having related weight & number value of input is assigned random value during beginning of learning process of DNN model. It is need to perform calculation by neural network into combined forms of input and weight data & the same is spread throughout deep neural model until final outcome is made.

It is also need to explain, the result is a function that maps output, inputs and calculations indicate a feature. Value of weights for our network needs to improve. It is required to measure various functions through the calculations used by network, according to different layers. Apart from calculation, all variable names used in our project must be described with accuracy.

- i. In this work, variable X is the input layer, by which we feed network data.

- ii. In this work, variable Y will be the required output which is found at end of DNN with operating its calculation to make end result equivalent to input X.
- iii. In this project, variable \widehat{Y}_h is reflecting our prediction & for getting out, network need to feed with input X. one of the variable Y in our project is ideal output which is created by network when data is feed to it.
- iv. One of important variable W in this work is weights of network layer.

A:- The adopted calculation: W X is product between (W and X), our hidden level first layer. It generates a weighted amount:

- i. Units of each layer are connected to each unit of the previous layer.
- ii. A weighted value is available for each of its network links.
- iii. To find out new value of each neuron which is sum of result of previous unit & current unit has been generated by multiply the previous unit.

It is also need to train network for increasing it's ability to learn mystery functions quicker & corresponding weight in every link has shown strength or weakness among them. Network should have flexibility to calculate linear calculation to move one point to other or vice-versa.

B: - there is linear equation: $X + b$, which is using in this work to relate relationship between input and output & this has been find out by product and sum. One of variable Z used to represent the output of function & equation to find out calculation.

$$Z1 = W1X + b1$$

Therefore; computing on every unit of each layer using calculation equation generate result of individual unit. Now it's need to implementation of algorithm using vectorized form in computer programing. This shows that matrix form is used to perform to combine all calculation of mathematical operation in every layer.

Gradient-Descent – In this proposed work, a gradient measure how much the output of a function changes if you change the inputs a little bit. Gradient of a function is often called its derivative & when gradients (derivatives) become smaller than its output become flat. Gradient with back-propagation algorithm are used particularly in deep learning, to determine weight changes implication of network parameter impact network output. One of the impotent charterer stick need to determine that vanishing gradients is become very small or empty then it is difficult to judge performance of system changes in this condition. Here it is also possible opposite issue that gradients may be burst. In this condition, network can become very unstable due to its values become very high. The various types of activation function are used for different benefits. But these functions can suffer from burst gradient issue. The values of variables W & b used in equation are not known in all variables. In this situation network must learn correct value of unknown variables so that correctly function is measured. To get the right W_1, b_1, W_2 and b_2 values, our network must be train & values of these variables must be initialize before we can start the training.

In this works, Relu and Sigmoid activation functions are used and forward function carries out the above calculations. Functional activities of deep neural network: -to find out variable Z1 then adding the first bias matrix b_1 , & system is multiply weights of the first layer to generate Z_1 . Using Z_1 , to generate A_1 with the Relu function. At the next step, it is needed to multiply weight matrix of second layer, A_1 (the system use first layer outcome as second layer 'input), & after that we need to perform addition of add second bias matrix to calculate, b_2 , so that Z_2 is generated. Sigmoid activation function is applied on Z_2 , for generate the A_2 , which is \widehat{Y}_h , the network output. At last, we ought to achieve: \widehat{Y}_h . and Y. After said activities, it requires to do addition a final network feature, the loss feature, to measure out this.

In this work, calculus and chain rule of derivatives are used, to understanding change in a variable when modifying a different variable. It is also need to explain the Chain Rule; it is the method of fine-tuning the weights of a neural network based on the error rate obtained in the previous iteration.

Now, author is need to discussed one of main important Back-Propagation algorithm [20] in ours work. It has been used to estimate changes in network's various parameters impact on its final loss. To understand the action chain law, select one of our parameters and starts with the loss equation:

$$Loss = -(Y \log Y_h + (1-Y) \log (1-Y_h))$$

In this equation, one parameter loss changes is required to calculate when a shift in Y_h and outcome. Chain derivatives has been continue until we reach W_1 , & derivative equation of the loss function to find out this derivative, is mentioned here.

$$dLoss_{Yh} = - (Y/Y_h - (1-Y)/(1-Y_h)) \cdot (Y_h - 1 - Y_h)$$

$$Y_h = \text{sigmoid}(Z2)$$

$$d\text{Sigmoid} = \text{sigmoid}(x) * (1.0 - \text{sigmoid}(x))$$

Hence, representing this difference equation as $d\text{Sigmoid}$, & in the form of easy writing. Hence:

$$dY_h_{Z2} = d\text{Sigmoid}(z2)$$

As explored, these two derivatives can already be chained (multiplied) for Z2 derivative.

$$dLoss_{Z2} = dLoss_{Yh} * d\text{Sigmoid}(z2)$$

$$Z2 = A1W2 + b2$$

4. RESULTS AND DISCUSSION

The complete work has been done in version-3 programming language of Python using its various libraries like pandas, scikit-learn, mat lab plots, etc. Figures and tables mentioned below in this paper exhibit the outcome of our implemented model which is equated with available approach like as K-NN, NB, RF and NN.

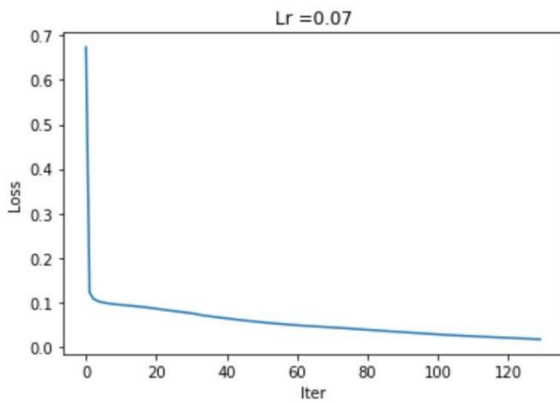


Figure-2 Reveal Computed Loss

Figure-2 indicates the loss graph; which reflects that computed loss minimum in this finding.

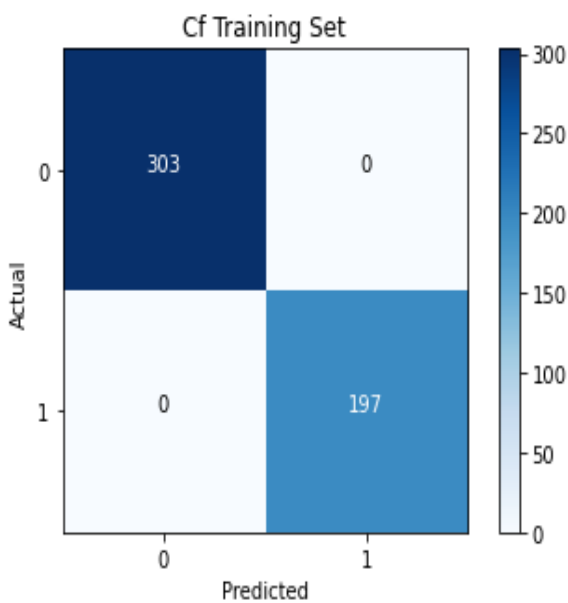


Figure-3 Confusion Matrix of Training Set

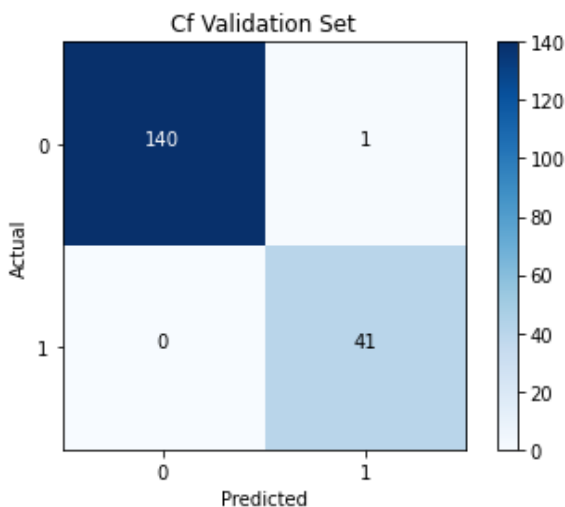


Figure-4 Confusion Matrix of Testing Set

The Wisconsin cancer dataset is divided into training and testing dataset in which out of total records, five hundred is

considered for training dataset & one hundred eight two for testing set. In this paper, confusion matrix of training set has been represented by figure-3 and figure-4 has exhibited confusion matrix for testing set. The training & testing data set accuracy are 100% & 99%.

Table-2 Comparison of existing algorithms with proposed classifier

Model	Accuracy (%)
KNN (K-Nearest Neighbor)	97.51
NB (Naïve Bayes)	96.19
Random Forests	96.85
Neural Networks	93.02
Deep Neural Network (Proposed)	99.45

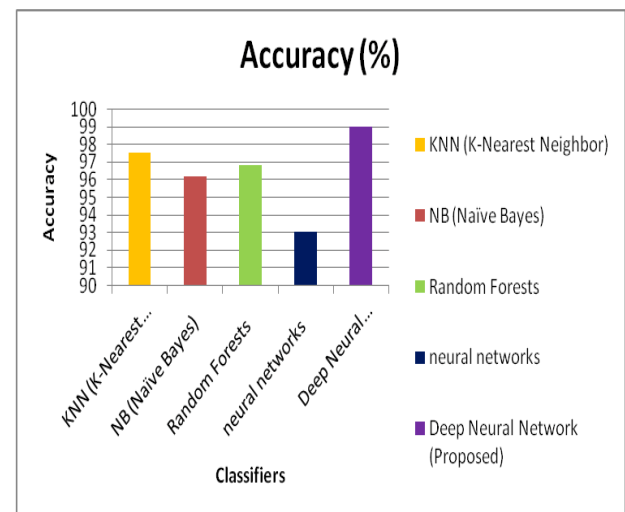


Figure-5 Accuracy Comparison Graph

Figure 5 has shown graphical representation of previously available classifier with the implement classifier. It has been found out that accuracy more than 99% which is better than other states of the art algorithms has been achieved by the proposed model.

5. CONCLUSION

Cancer is a critical disease, it has been affected millions of people across the world and people are scared due to its fatal nature. It is found in many cases that disease can be cured if diagnosed at early stage with proper medical follow up and attention. The author used deep learning which is artificial intelligence techniques, where thought processes of a human brain are mimic by machine. This paper proposed a two-layer deep neural network model having gradient descent & back propagation algorithm prediction is excellent to classified breast cancer on Wisconsin Breast Cancer dataset & outcome of the result shows that deep neural network having 2-layer is precise (99.45%) with minimum error as compared to NB classifier (96.19%), NN classifier (93.02%), and KNN classifier (97.51%). In the future, this work is extended for

cancer classification on image of MRI, X-ray, CT-Scan, database, etc.

6. REFERENCES

- [1] A. McGuire, J. A. L. Brown, and M. J. Kerin, "Metastatic breast cancer: the potential of miRNA for diagnosis and treatment monitoring," *Cancer Metastasis Rev.*, 2015.
- [2] M. A. Schonberg, E. R. Marcantonio, D. Li, R. A. Silliman, L. Ngo, and E. P. McCarthy, "Breast cancer among the oldest old: Tumor characteristics, treatment choices, and survival," *J. Clin. Oncol.*, 2010.
- [3] M. D. Ganggayah, N. A. Taib, Y. C. Har, P. Lio, and S. K. Dhillon, "Predicting factors for survival of breast cancer patients using machine learning techniques," *BMC Med. Inform. Decis. Mak.*, 2019.
- [4] S. Horibata, T. V. Vo, V. Subramanian, P. R. Thompson, and S. A. Coonrod, "Utilization of the soft agar colony formation assay to identify inhibitors of tumorigenicity in breast cancer cells," *J. Vis. Exp.*, 2015.
- [5] R. Turkki *et al.*, "Breast cancer outcome prediction with tumour tissue images and machine learning," *Breast Cancer Res. Treat.*, 2019.
- [6] A. Reeves and Y. Xie, "Software for density calculation and timing of surgery in subsolid nodules," *J. Thorac. Oncol.*, 2015.
- [7] R. Nair and A. Bhagat, "Feature selection method to improve the accuracy of classification algorithm," *Int. J. Innov. Technol. Explor. Eng.*, 2019.
- [8] A. M. Abdel-Zaher and A. M. Eldeib, "Breast cancer classification using deep belief networks," *Expert Syst. Appl.*, 2016.
- [9] E. Rolls and A. Treves, *Neural Networks and Brain Function*. 2012.
- [10] P. Subramanian, N. O. Oranye, A. M. Masri, N. A. Taib, and N. Ahmad, "Breast cancer knowledge and screening behaviour among women with a positive family history: A cross sectional study," *Asian Pacific J. Cancer Prev.*, 2013.
- [11] R. Nair and A. Bhagat, "Genes expression classification using improved deep learning method," *Int. J. Emerg. Technol.*, 2019.
- [12] K. A. Parato, D. Senger, P. A. J. Forsyth, and J. C. Bell, "Recent progress in the battle between oncolytic viruses and tumours," *Nature Reviews Cancer*. 2005.
- [13] Y. Li, S. Li, X. Meng, R. Y. Gan, J. J. Zhang, and H. Bin Li, "Dietary natural products for prevention and treatment of breast cancer," *Nutrients*. 2017.
- [14] S. Alelyani, J. Tang, and H. Liu, "Feature Selection for Clustering: A Review," in *Data Clustering*, 2019.
- [15] S. Bashir, U. Qamar, and F. H. Khan, "Heterogeneous classifiers fusion for dynamic breast cancer diagnosis using weighted vote based ensemble," *Qual. Quant.*, 2015.
- [16] H.K. Timmana and Rajabhushanam C, "Breast malignant detection using Deep Learning Model," *IEEE.*, 2020.
- [17] V. Chaurasia and S. Pal, "Data mining techniques: To predict and resolve breast cancer survivability," *Int. J. Comput. Sci. Mob. Comput. IJCSMC*, 2017.
- [18] Jing Zheng¹, Denan Lin², Zhongjun Gao³, Shuang Wang⁴, Mingjie He⁵, Jipeng Fan⁶, "Efficient AdaBoost Algorithm for Breast Cancer Detection and Early Diagnosis," in *IEEE Proceedings*, 2017.
- [19] N.Pradhan and V.S. Dhaka, "Comparative Analysis to Predict Breast Cancer using Machine Learning Algorithms," *IEEE.*, 2020.
- [20] M. Cilimkovic, "Neural Networks and Back Propagation Algorithm," *Fett.Tu-Sofia.Bg*, 2010.