Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 10, 2025

# SYMPTOM-BASED BREAST CANCER CLASSIFICATION USING DEEP NEUTRAL NETWORK

**Sumia Kanwal[1], Naveed Ahmad[2], Zaid Sarfraz[3], Muhammad Talha Jahangir[*4], Muhammad Rehman[5], Muhammad Hassan[6], Anas Mahmood[7]**

[1]*Department of Computer Science, University of southern Punjab Multan, Pakistan*
[2]*Department of Computer Science NCBA&E Multan Campus, Multan, Pakistan*
[3,*4,5,6,7]*Department of Computer Science, MNS-University of Engineering & Technology, Multan, Pakistan*

[*4]mtalhajahangir@mnsuet.edu.pk

DOI:https://doi.org/10.5281/zenodo.17375155

**Keywords**
*Breast Cancer Detection, Deep Neural Networks, Malignant Tumor, Benign Tumor, Machine Learning*

**Article History**
*Received: 27 August 2025*
*Accepted: 05 October 2025*
*Published: 17 October 2025*


**Corresponding Author: ***
**Muhammad Talha Jahangir**

**Abstract**
*Breast cancer is among the commonest cancer in the world, as the incidence rates have been on the rise and have posed a big burden to the healthcare systems. To increase the chance of survival and decrease the mortality, it is important to detect and diagnose the problems early and properly. Recent literature has applied machine learning and deep learning to the classification of breast cancer, but numerous models, such as Support Vector Machines (SVM) and Random Forests, have been challenged by lack of accuracy and generalization. We used a Deep Neural Network (DNN) to determine benign and malignant tumors using the Wisconsin Diagnostic Breast Cancer (WDBC) dataset in this study. The data has five hundred six hundred and ninety-nine samples and thirty features which were obtained as nuclei of breast mass cells, the data has undergone pre-processing by eliminating irrelevant data and has been standardized to achieve a better model performance. In order to avoid overfitting, the DNN architecture was trained until 50 epochs and early stopping was implemented after 36 epochs. The model was successful at 98.63 and 97.83 in training and validation accuracy respectively, and it can be inferred that the model is performing well in the classification of breast tumors. The findings suggest that the developed DNN model is a viable instrument in the early diagnosis of breast cancer, which is a crucial issue in the prior machine learning models*

## INTRODUCTION

Cancer is considered to be one of the most complex and dangerous diseases in the current century and it kills millions of people every year, as well as affects the families, businesses, and health care systems the world over massively. Regardless of social or ethnic distinctions, it impacts people on every continent. Because of its high incidence in women and the complex difficulties in diagnosing and treating it, breast cancer is the most concerning of all its types [21]. The development of a malignant lump results from the aberrant division of normal cells in the breast's epithelial tissue. Because of its aggressive nature and propensity to spread to other organs, breast cancer is one of the world's top causes of death for women.Uncontrolled cellular proliferation beyond the body's regulatory processes leads in a malignant tumor, which is categorized as breast cancer when it originates in breast tissue [22]. Cell division, growth, and death occur in a systematic manner under normal circumstances. Cells may divide excessively as a result of hormonal imbalances, environmental toxins, or abnormalities in the DNA of the cell. Both benign and malignant cancers can result from this aberrant growth. Malignant tumors have the capacity to penetrate surrounding tissue and travel to distant organs via the lymphatic or

circulatory systems, while benign tumors often stay isolated and do not invade adjacent tissues [23].

The fact that breast cancer can progress silently with minimal to complete symptoms makes the task of identifying it at its early stages one of the greatest challenges to medical professionals [24]. Nevertheless, early diagnosis can greatly improve survival rates by giving patients a chance to intervene on time and apply the appropriate treatment modalities [25]. Even though they were widely used as detecting methods, their techniques such as tissue biopsy, MRI, ultrasound and mammography continue to experience problems in sensitivity and specificity. Thus, to find methods of more efficient and accurate detection they have turned to computational intelligence.

Integrating deep learning (DL) and machine learning (ML) models has been highly promising in recent years to enhance the early detection and classification of breast cancer. These advanced computers are able to scan large amounts of data including genetic data, biopsy slides as well as mammography images, to identify patterns that could not be spotted by the naked eye. At that, convolutional neural networks (CNNs) have become the most impressive deep learning models that have been proven to perform at a very high level of recognizing malignant lesions and distinguishing between different stages of cancer. Notwithstanding these developments, a number of issues still affect neural network performance, including class imbalance, model interpretability, and hyperparameter tuning. Furthermore, because of poor model construction, a lack of data, and a restricted capacity for generalization, previous research frequently had trouble with multi-class classification problems.

Breast cancer is a fast growing sector of the medical business from the standpoint of the economy and healthcare market. The global market value of chemotherapy and related therapies is predicted to increase significantly, from 18.8 trillion USD in 2022 to around 43.2 trillion USD by 2032, as shown in Figure 1, according to the Breast Cancer Treatment Market Categorization (2022–2032) study. Numerous reasons have contributed to this exponential expansion, including the rise in breast cancer incidence worldwide, increased awareness of early diagnosis, increased investment in cutting-edge diagnostic technology, and the creation of tailored medicines like precision oncology and immunotherapy. The use of AI-assisted technologies in clinical settings and improved access to healthcare services in developing nations have also contributed to the market expansion for breast cancer treatments.

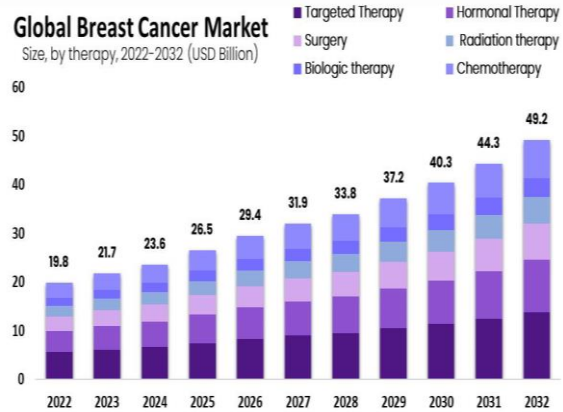


**Fig. 1.Global Trend of Breast Cancer market [20]**

The WDBC dataset, which includes 569 samples that were categorized as either benign or malignant based on 30 features extracted from the nucleus of breast mass cells, was used in this investigation. By preprocessing the data and splitting it into test sets and training using stratification, we were able to guarantee class parity. To ensure a mean of unity and unit variance, the attributes were standardized using a conventional Scaler to maximize the input for training the model. A deep neural network, or DNN, was constructed with two hidden layers with 64 and 32 neurons each that employed ReLU activation in order to prevent overfitting. Additionally, a dropout layer with a rate of 0.3 was added. A sigmoid active function, designed

especially for binary classification, was employed in the output layer. The model was built using binary cross-entropy loss and the Adam optimizer with 0.001 learning rate. Early termination of training was used when the loss on validation stopped improving after 36 epochs. The findings indicated that the model was very accurate in distinguishing benign and malignant tumors, and this demonstrates how the model can be used to diagnose breast cancer at an early stage. The primary distinction of the study given is that it has an optimized DNN architecture and standardized pre-processing data, which are granted to be effective in generalization and stable in performance without overfitting. The outline of this paper is as follows: Section II is the literature review. The architecture of the deep neural network has been described in Section III, the datasets used have been described in Section IV and the examples of the entire process have been provided in Section V. We outline the evaluation measures in Section VI, and we show the results of the experiment in Section VII. The conclusion of section VIII is an overview of the entire work and future initiatives.

**Literature Review**

The use of machine learning (ML) and deep learning (DL) models in breast cancer detection has grown very quickly over the past decade. The hybrid models with multiple feature selection, classification schemes, and hybrid frameworks have been proposed and optimized in various studies. Each of the proposed methods contributes uniquely to the increase of computational and diagnostic accuracy and the external validity of different datasets. Despite these developments, several significant challenges to the research on breast cancer prediction still exist, the most notable of which are the lack of data between benign and malignant cases, the small sample size, the absence of data diversity, and noisy or missing data which complicate the task of assessing and training the model. Also, the interpretability of machine learning predictions remains largely unstudied, especially in the clinical setting where model transparency is equally crucial to model accuracy.

To categorize breast cancer, the research in [6] performed a comprehensive comparison of the ensemble machine learning and the classical approach. The authors additionally employed ensemble methods like Random Forest (RF), XGBoost, and AdaBoost besides Support Vector Machine (SVM), Naive Bayes (NB) and K-Nearest Neighbors (KNN). XGBoost was able to outperform the other models due to its gradient boosting optimization and its ability to address non-linear interactions among a high number of dimensions. This finding underlines the increase in the importance of boosting-based models to structure biological data.

In [7], eleven machine learning algorithms have been considered with the care of arriving at the most reliable dependable classifier that can predict breast cancer. The models were contrasted with regards to a set of performance measures: accuracy, precision, recall, F1-score, and specificity. The study demonstrated optimal mixed results in additional Randomized Trees in every measure. In order to make the classification more robust to the classifier and decrease overfitting, the authors insisted on the fact that it features randomization and ensemble averaging matter.

The identical form of study that also attempted to develop a reasonable diagnostic model with the help of ML was mentioned in [8] and utilized information that is found in the Wisconsin Diagnostic Breast Cancer (WDBC). The performance of the detection as assessed by the authors of the research was evaluated based on the Random Forest (RF), Decision tree (DT), Logistic Regression (LR), and KNN models. It was found out that predictive capacity of any classifier was significantly enhanced due to feature scaling and normalization, and data preparation is a significant factor of consideration when it concerns the model performance.

The Decision Tree and the Logistic Regression models were chosen in particular for comparison between them in [9]. The main aspect of research was preprocessing algorithms (dimensionality reduction, and elimination of outliers). The Decision Tree model was better understood by clinical understanding, and the logistic regression was better understood by linearly separable data while both models have acceptable performances.

This approach incorporated four advanced ensemble classifiers, namely Extreme Gradient Boosting (XGBoost), AdaBoost, Random Forest (RF) and Logistic Regression (LR) over an already processed dataset in a study by [2]. After a rigorous testing procedure, the researchers

concluded that the regularization approach of XGBoost, which minimizes overfitting and stabilizes training, was the factor that made this algorithm receive the most classification accuracy.

The authors of [10] accept and test five machine learning techniques, including SVM, LR, KNN, DT, and RF on the Wisconsin Breast Cancer (WBC) dataset and concluded that SVM outperforms all those, and the high accuracy of 95.01 was achieved. The Benign versus Malignant classification was how SVM was able to handle overlapping features in its classification between benign and malignant data, as it is known to have maximized a margin, which showed that it was applicable to data acquired by medical images.

In order to facilitate early detection, [11] compare several datasets, namely WDBC, Fmc, and SEER. Based on their results, the highest-ranking classification results on SEER dataset were produced by Decision Tree (DT) and the best classification results on Women Health dataset were obtained by RF and SVM. It demonstrates the importance of a dependency on datasets and that the behavior of an algorithm can be influenced by the properties of datasets including feature correlations and class distribution.

The investigation in [12] used the WBC dataset that comprises six morphological and textural characteristics linked to the tumor characteristics. The scientists applied SVM, KNN, RF and NB in order to detect malignancy. They demonstrated that the combination of hybrid features helps to increase the stability and interpretability of the models. Similarly, [13] also employed correlation-based feature selection (CFS) to select ten key features out of an original thirty features to determine the potential of the WDBC dataset to diagnose. Having considered twelve machine learning models, the researchers established that ensemble-based classifiers, namely, RF and XGBoost, yielded the best reliable predictions.

A follow-up comparison study in [14] was the study of the performance of RF and KNN classifiers at three different train-test split ratios (7030, 8020 and 9010). The results showed that RF performed better with less training data, and its overall generalization and its accuracy were better at all ratios. In a study that places emphasis on feature-engineering [15], apriori algorithms (AR) and association rule mining were used in enhancing feature selection. The study proceeded to apply SVM+AR as a combination classifier resulting in less computing costs and high classification accuracy. This hybrid approach demonstrated how the combination of data mining and machine learning methods can enhance interpretability and predictive power.

In [16], a hybrid CNNGRU (Convolutional Neural NetworkGated Recurrent Unit) model was suggested to conduct image-based diagnosis based on PCAM dataset that examined the application of deep learning to breast cancer analysis. Having an accuracy of 86.21% and an F1-score of 86 the model was highly promising in the recognition of histopathology images. The hybrid method was superior in offering a generalization compared to solo CNNs because it used the sequential memory capabilities of GRU and the spatial feature extraction of CNN.

Lastly, in order to differentiate between benign and malignant cases, [17] combined multiple machine learning classifiers using WBC dataset. In order to the better interpretability and the better model accuracy, the authors put a large emphasis on the feature selection and the inter-feature correlation analysis. The results indicated that Random Forest (RF) and Support Vector Machines (SVM) yielded the highest accuracy rate of (96.5%) and are reliable frameworks for AI-driven and automated breast cancer diagnostic systems. The result of the study found that compared to the shallow model, ensemble-based classifiers and kernel-based classifiers have a better performance in processing the medical data set with high dimensions.

**Deep Neural Network Architecture**

A Deep Neural Network (DNN) is one of the most powerful computational models in artificial intelligence that is inspired by the structure and function of the human brain. The principal feature that differentiates it from the traditional Artificial Neural Network (ANN) is the depth, that is, the presence of several hidden layers between the input and output layers. As a DNN has increasing depth, it is able to capture much more abstract and hierarchical data representations and can learn complex correlations that are often out of the reach of shallow networks to accurately predict.

Raw data to the learning system passes through first layer of DNN. It can make use of sensor data, images, time-series data, numerical properties of data, and other written data. The whole set of

input neurons corresponds to a single dataset feature; hence the number of input neurons give the dimensionality of the issue space. Apart from the bias mitigation caused by irregular feature scales, proper preprocessing increases the speed of convergence during training.

One or more hidden layers follow the input layer in a DNN and they are responsible for learning hierarchical representations and feature extraction. The various artificial neurons that constitute each hidden layer perform a non-linear activation function upon performing a weighted summation of the inputs. As the data is transmitted through these layers, the network is progressively trained to convert simple input features into conceptually higher ones. In image processing, the simple colours and edges identified in the first hidden layer may be more complex in the second layer, which in turn may be more complex in the third layer, a form, a texture, and finally an entire object or even an anatomical system.

Activation is like the Sigmoid, Hyperbolic Tangent (tanh) and the Rectified Linear Unit (ReLU) which allow the hidden layers to represent non-linear interactions. The most popular of them in modern systems is ReLU due to its computational efficiency and the ability to overcome the vanishing gradient problem. ReLU can be described as a non-linear redefinition of all negatives to 0 and the positives will pass without being altered. This ensures neuronal selectivity of activation leading to sparse representations which optimize generalization and performance.

The DNN has trainable parameters which are the weights and biases of the connections among the neurons. Once the previous layer has provided inputs to it, each neuron combines a bias term, multiplies the inputs by the corresponding weights, and applies the output to an activation function. During the training period, these parameters are adapted repeatedly to reduce the loss function which is used to measure the difference between the actual target and the expected output. To perform this correction process, the backpropagation algorithm, which uses the chain rule of calculus in reverse, i.e. backward, on the loss function to determine gradient of the loss function with respect to each weight and bias in the loss term is used.

These parameters are effectively updated by optimization techniques such as Mini-Batch Gradient Descent, Stochastic Gradient Descent (SGD) and more advanced techniques such as Adam and RMSprop. To achieve better convergence stability and reduce oscillations in the course of training, some solutions incorporate momentum and adaptive learning rate schedule.

Another essential concept in DNNs is regularization because it prevents overfitting, which is a common issue in the case of a network committing to memory training data instead of general trends. Regularization techniques such as L1 (Lasso) and L2 (Ridge) penalty term are often added to the loss to restrict the size of the weights. Another widely used approach is the so-called batch normalization, which helps stabilize learning dynamics and allows us to use larger learning rates by normalizing between-minibatch intermediate activations.

The last classification or prediction result is generated by the output layer of a DNN. The structure of it relies on the type of the problem: in binary classification, a single neuron with a sigmoid activation is typically used to produce probabilities between 0 and 1; in multi-class classification, a softmax layer is used to produce the normalized probabilities across all the possible classes; and in regression, a linear activation function is used to produce continuous numerical predictions.

Depending on the choice of activation functions, the number of hidden layers, the number of neurons per layer, and the learning rate, are some of the hyperparameters that define the overall structure of a DNN. To optimize the performance, the appropriate hyperparameters should be selected and most of the time this is facilitated by grid search, random search, or Bayesian optimization. Moreover, the advent of deep learning systems such as TensorFlow, PyTorch, and Keras has enabled deep learning to be applied to a wide spectrum of study fields due to the greatly simplified process of constructing, training, and evaluating complex DNN and their structures.

DNNs have been successfully applied in many applications in the real world, such as speech recognition, natural language processing, medical picture analysis, and biomedical signal categorization. Since it is possible to achieve that state through the automatic identification of discriminating features in the raw medical image or patient information without any significant

manual feature engineering, DNNs have a high success in the breast cancer screening scenario. Figure 2 [19] represents the deep neural network architecture. The deep neural network architecture is presented in Figure 2 [19].

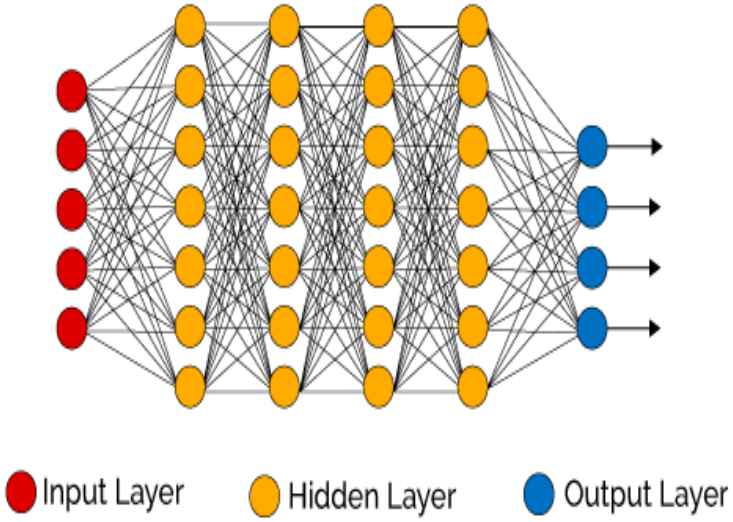


**Fig. 2.Deep neural network architecture [19]**

**Dataset**

**Dataset Description**

The data in this paper was acquired in Comma-Separated Value (csv) format and was taken out of the publicly accessible Kaggle listings [18]. The Wisconsin Breast Cancer Diagnostic (WBCD) dataset, originally designed and in the State of Illinois Clinic in Madison, Wisconsin, USA, was utilized in this study. This dataset has become one of the most popular and the most reliable standards in the field of medical image analysis and cancer diagnostic research. It gives crucial diagnostic data with the help of which benign and malignant breast cancers are distinguished by taking the numerical characteristics of the microscopic pictures of breast tissues.

The dataset is arranged as 32 columns, each of them is unique and performs a unique role in classification, and prediction activities. In order to have every data instance that can be easily recognized, the first column of the dataset would have the unique identification (ID) that is associated with any given patient record. The second column displays the diagnostic outcome of the tumor being benign or malignant. In this column, a binary system of classification is used, with a 0 implying benign and 1 implying malignant tumor.

The results of quantitative measurements of the digital analysis of data on the breast cell nuclei are presented in the rest of the 30 columns located between the third and the thirty-second. These attributes are the meaning, the standard error of the mean, and the worst or largest value of the meaning of eleven important cellular properties. All these characteristics define the morphological architecture, texture, and geometry of the nucleus of cancer cells. The features are largely based on the microscopic examination of the biopsy samples acquired with the use of Fine Needle Aspiration (FNA), a minimally invasive procedure commonly used to collect breast tissue used to detect the presence of cancer. A microscope is used to analyze every sample in a pathology laboratory.

To guarantee high order of precision and accuracy in the calculation of features, the numerical representations in each of the observations in the dataset have no less than four significant digits. The dataset has not any null or missing values, which enhances its applicability and reliability when used in data-driven modeling methods. The property allows researchers to operate on these complex classification models without the need to do to data-cleaning or imputation work.

The dataset contains 10 real-valued characteristics, which are a detailed description of the morphological properties of breast cells, such as their radius, texture, perimeter, area, smoothness, compactness, concavity, symmetry and fractal dimension. These parameters

statistical measures such as mean, standard error and extreme values help the computer to learn complex patterns associated with malignant changes. These parameters are important indices of abnormality of cells. Due to the comprehensive quantitative character of the dataset, even smaller alterations in cell structure signaling the onset or dissemination of breast cancer can be detected by the use of computational models.

Altogether, the WBCD dataset is an example of a dataset that can be used to evaluate the efficacy and reliability of deep learning and machine learning algorithms in medical diagnostics. It can be applied to experimental analysis and model validation on the spot due to its clinical observational history, well-designed structure, and missing data. The dataset represents a crucial resource to the researcher seeking to enhance automated cancer detectors and early diagnostic technologies due to its accuracy and comprehensiveness, which tremendously enhance the accuracy of breast cancer forecasts [18].

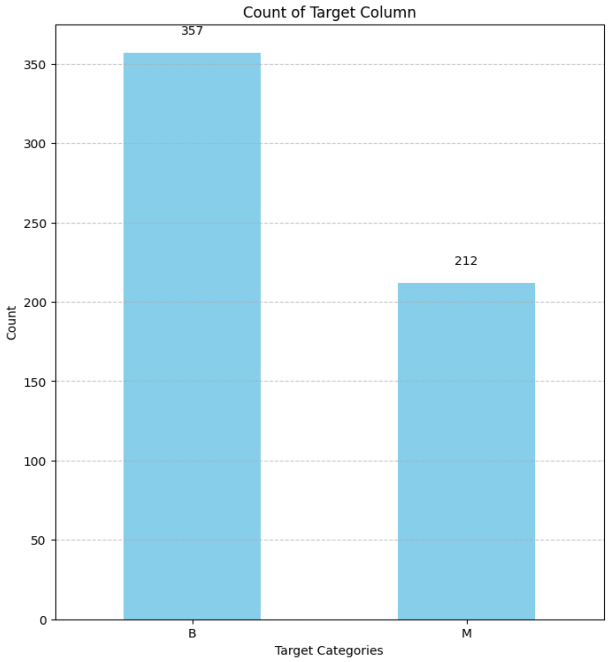


**Fig. 3.Class Distribution**

The 569 samples included in the dataset used in this study are all individual observations of the breast tissues of the Wisconsin Breast Cancer Diagnostic (WBCD) dataset [18]. The imbalance in classes in this dataset is also quite apparent and this aspect significantly affects the capability and fairness of classification systems. The benign type (abbreviated B) is the most specific class that consists of 357 individual examples. The malignant group (labeled as M) is relatively underrepresented in comparison to the benign one, only 212 samples. This disproportion in the two groups is reflected in figure 3 that depicts the difference in the number of observations of the two classes.

Such a biased distribution of classes inhibits the formation of reliable prediction models considerably. Machine learning and deep learning algorithms, in particular, the ones that are trained through supervised learning, are often biased to the majority class. It implies that a model that is trained on an imbalanced dataset can be highly accurate in general, but it might also fail to identify the minority category, which in this instance is the cancerous or the malignant cases. In medical diagnosis, where the accurate distinction of malignant tumors is far more critical than the correct classification of benign specimens, this so-called phenomenon or class bias or majority dominance can be highly detrimental.

Thus, it is important to correct this imbalance to ensure the robustness and clinical relevance of the trained model. To reduce the effects of imbalance in the classes, several data pretreatment and resampling methods are typically considered. Some of them, such as cost-sensitive learning, random oversampling, and the

Synthetic Minority Oversampling Technique (SMOTE) are often used to enhance the representation of the minority class. The model is encouraged to be equally treated to both classes in the training stage by synthetically creating instances or adjusting weights of the classes, that increases the sensitivity and generalization ability of the model to malignant cases.

**Table 1 described the details of each feature used in WBCD dataset.**

**TABLE I. FEATURE DESCRIPTION**

| Feature Name | Feature Description |
|---|---|
| Radius | Measures the distance from the tumor's center to its edge. |
| Texture | Quantifies the variability in gray-level intensities within the tumor. |
| Perimeter | Length of the tumor's boundary. |
| Area | Number of pixels contained within the tumor. |
| Smoothness | Assesses the variation in radius lengths around the tumor's boundary. |
| compactness | Perimeter2/Area |
| Concavity | Evaluates the depth of concave regions along the tumor's boundary. |
| Concave points | Counts the number of concave regions in the tumor's contour. |
| symmetry | Measures the similarity between the two halves of the tumor. |
| Fractal dimension | Represents the complexity of the tumor's boundary. |

**Methodology**

As shown in Figure 4, our work makes use of the Breast Cancer dataset, which has been preprocessed to guarantee the best possible performance during the training of the model. In order to ensure that the training data was adequately validated, a 10% validation split was used. While training, this was useful for tracking how well the model generalized to new data. The model was trained using the other 90% of the data. In order to categories breast cancer, we trained a neural network with deep learning (DNN). Ideal for managing complicated data transformations, the model architecture comprises of several dense layers that are completely coupled and use ReLU activation functions. By adding batch normalization layers after each thick layer, an enormous boost was achieved. By standardizing the result of the preceding layer, this approach helps stabilize and speed up the training processes. In order to avoid overfitting, we also included layer of dropout with a 30 percent dropout rate. The dropout layers improve the model's generalizability by randomly deactivating a percentage of neurons during each training cycle. A the sigmoid activate neuron makes up the model's last layer; this kind of neuron is well-suited to binary classification since it produces a probability score (from 0 to 1) that represents the likelihood of a certain class. (benign or malignant).

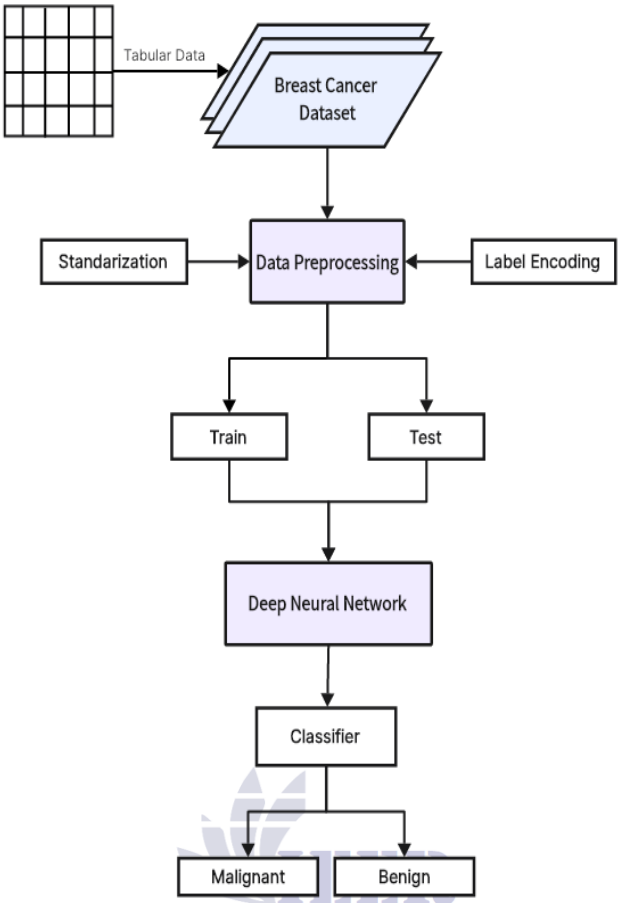


**Fig. 4.Flowchart of the proposed methodology**

**TABLE II.**
SUMMARIZING THE ARCHITECTURE OF PROPOSED DEEP NEURAL NETWORK (DNN) MODEL

| Layer (Type) | Output Shape | Number of Parameters |
|---|---|---|
| Flatten (Input) | (None, 30) | 0 |
| dense (Dense) | (None, 64) | 1,984 |
| batch_normalization | (None, 64) | 256 |
| dropout | (None, 64) | 0 |
| dense_1 (Dense) | (None, 32) | 2,080 |
| batch_normalization | (None, 32) | 128 |
| dropout_1(Dropout) | (None, 32) | 0 |
| dense_2 (Dense) | (None, 16) | 528 |
| dense_3 (Dense) | (None, 1) | 17 |

The Adam optimizer, an optimization method with flexible learning rate characteristics, was utilized to construct the model. After conducting early trials, it was determined that the rate of learning of 0.001 would provide the best balance between velocity of convergence and model performance. Binary cross-entropy, a common loss function for binary classification problems, efficiently evaluates the discrepancy between the actual labels and the projected probabilities. Consequently, it was chosen for this particular challenge. Early halting was used to avoid overfitting. In the event that the model's performance fails to change within five consecutive epochs, this callback will terminate the training process based on the validation loss.

To avoid being inaccurate to the training data, the model uses the weights in the epoch that performed the best during validation. Batch size was used for the model's training. of 32 over a maximum of 50 epochs. The summary of proposed deep neural architecture is given in Table 2.

**Data-Preprocessing**

**Cleaning**

In our study, the preprocessing of the Wisconsin Diagnostic Breast Cancer dataset was an essential step to ensure the data was clean and ready for model training. First, the dataset contained an 'id' column that was irrelevant to the classification task, so it was removed to prevent any influence on the model. Then, we categorized the data based on the target variable, diagnosis, and computed the means of the features with each of the classes (benign and malignant). This gave me a summary of the difference between the two categories in each of the features.

**Standarization and Normalization**

We subsequently standardized the data with the StandardScaler which centers the data to a mean of 0 and a standard deviation of 1, which made the data the same across features. This normalization process is advantageous to machine learning models because any of the characteristics will be excessively dominant due to its magnitude. The distributions of the features of the cancerous and benign cases were compared in a bar plot after the normalized data were inserted in a newly created DataFrame.

**Label Encoding**

A systematic preparation of the dataset after the completion of the normalization process was done to ensure the maximum performance and compatibility with machine learning and deep learning algorithms. To avoid the features of higher numerical value ruling the learning process, the normalization phase was necessary in the scaling of the numerical features to a similar range. This phase ensured that all the features made their contribution during the weight updates and model optimization. The focus was then made on the goal variable, "diagnosis" which initially comprised categorical values of string, with the B denoting benign cases and the M denoting malignant cases. The LabelEncoder method of the Scikit-learn library was applied to convert the category target variable into a numerical image so that it can be processed with algorithms more easily. This encoding assigned the numerical value of 0 to benign cases and the number 1 to malignant cases. Besides simplifying computational processing, this binary translation enabled the models to interpret the diagnostic results in a way that would be suitable in performing classification tasks and compatible with mathematics.

To gain a better understanding of the behavior of normalized features and their role in differentiating between the two diagnostic categories, exploratory visualizations were carried out following encoding. A thorough visual depiction of the normalized feature values is shown in Figure 5, which offers important information about each feature's discriminative ability. The distribution range of normalized values across all characteristics is depicted by the horizontal bars in this image; benign samples are represented by blue bars, whereas malignant samples are represented by red bars. In some areas of the graph, the distinct color separation draws attention to traits that are highly predictive. Features like "concave points_worst," "perimeter_worst," and "concave points_mean" stand out among the qualities that were analyzed because they clearly distinguish between benign and malignant findings. This obvious difference implies that these characteristics have high discriminative strength and are essential for correctly identifying the malignancy of breast cancers. According to the separation, the cell nuclei of malignant tumors often display more irregular and concave structural patterns, which is in line with established pathological results in the literature on medical imaging. However, characteristics like "texture_mean" and "fractal_dimension_mean" exhibit a clear overlap between the two diagnostic groups, indicating that these characteristics would not be sufficient to distinguish between benign and malignant instances on their own.

The importance of feature engineering and feature selection in creating a high-performing classification model is highlighted by this image. The predicted accuracy of the model may be greatly increased and the chance of misclassification decreased by using characteristics with obvious separability. Additionally, by

comprehending feature distribution patterns, researchers may spot noisy or redundant traits that could create bias or lower model effectiveness. Therefore, dimensionality reduction techniques like Principal Component Analysis (PCA) or Recursive Feature Elimination (RFE) can be used to either alter, combine with other qualities, or remove features that show low discriminating potential before the final training step.

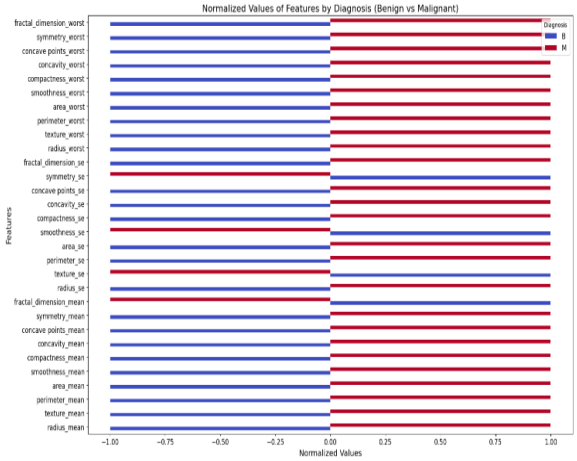


**Fig. 5.Normalized values of features by diagnosis (Benign vs Malignant)**

**Evaluation Metrics**

**Accuracy**

The accuracy of a diagnostic test or classifier is calculated as the number of true positive and true negative results divided by the total number of results. It can be expressed by (1).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (1)$$

**where, TP, TN, FP, and FN stand for true positive, true negative, false positive, and false negative respectively.**

**Precision**

Precision is the proportion of correct positive results, and is calculated by (2).

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

**Recall**

Recall represents the proportion of correctly identified positive cases out of all actual positive instances. It indicates the model's ability to accurately detect malignancies, as defined in equation (3).

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

**F1-Score**

The harmonic mean provides a balanced measure of precision and recall. It reflects the performance of deep learning models in terms of precision, recall, and F1-score, and is calculated using equation (4).

$$\text{F1Score} = 2 \times \frac{\text{Precision} \times \text{recall}}{\text{Precision} + \text{recall}} \quad (4)$$

**ROC**

A graph that illustrates the performance of a classification model across various thresholds. The ROC curve represents two key parameters: the True Positive Rate (TPR) and the False Positive Rate (FPR), as defined by equations (5) and (6).

$$TPR = \frac{TP}{TP + FN} \quad (5)$$

$$FPR = \frac{FP}{FP + TN} \quad (6)$$

**Experiment Setup and Results**

**Experiment Setup**

The experiments were conducted on a high-performance setup with the following specifications:
CPU: AMD RYZEN 9 5900X
GPU: NVIDIA GEFORCE RTX 4080 SUPER 16G VENTUS 3X OC
Memory: 32 GB RAM

Here, we examined the Deep Neural Network's (DNN) performance of the dataset after its deployment. There was a cap of 50 epochs for training, however because to convergence, it was terminated early at epoch 36. The validation accuracy and the final training accuracy were 97.83 and 98.63, respectively. A subsequent decrease in 0.0601 showed that pattern learning was effective and thereafter the loss of the model decreased. The model performed well with a test loss of only 0.0796 and accuracy of 97.37 per cent on the test set which had some unknown data. These results show that the model has an impressive generalization ability and that it can effectively distinguish between benign and malignant cases.

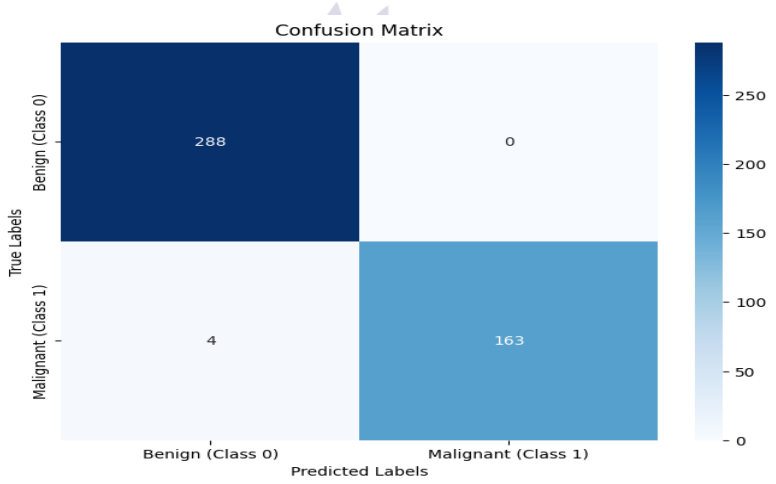


**Fig. 6.Confusion matrix**

The confusion matrix gives an overview of how the classification model performs its task of differentiating between benign (Class 0) and malignant (Class 1) cases. It means 163 malignant cases were correctly diagnosed as positive and 288 benign cases were correctly diagnosed as negative. It is worth mentioning that False Positives are zero, which means that no benign cases were mistaken as malignant. There are, however, four False Negatives which are malignant cases falsely ruled as benign.

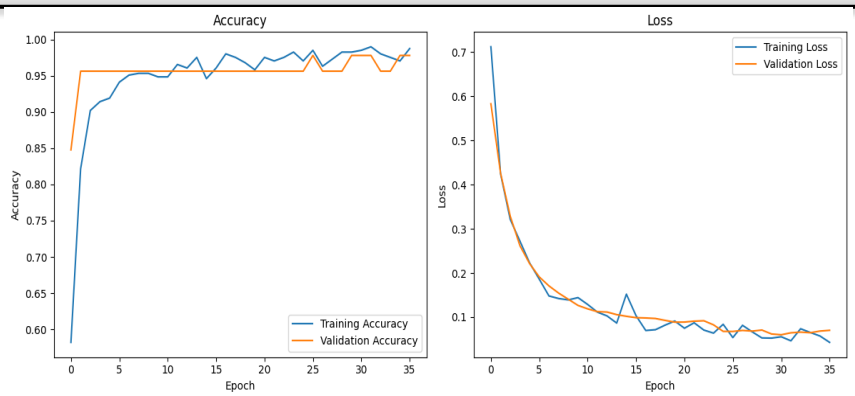


**Fig. 7.Performance graph of proposed model**

The model learning curves as shown in Fig. 7 indicate both accuracy and loss with the increasing training epochs. It is worth noting that the training and validation accuracy are steadily increasing with accuracy of around 98 percent, which implies effective learning of the training data as well as good generalization to unseen cases, thus, indicating a desirable lack of overfitting. Simultaneously, the loss graph reveals a corresponding decrease in both training and validation loss over the epochs, with the close alignment of the curves further reinforcing the model's ability to generalize effectively. Overall, these learning curves reflect the successful training of a model with strong predictive performance and reliable generalization capabilities.

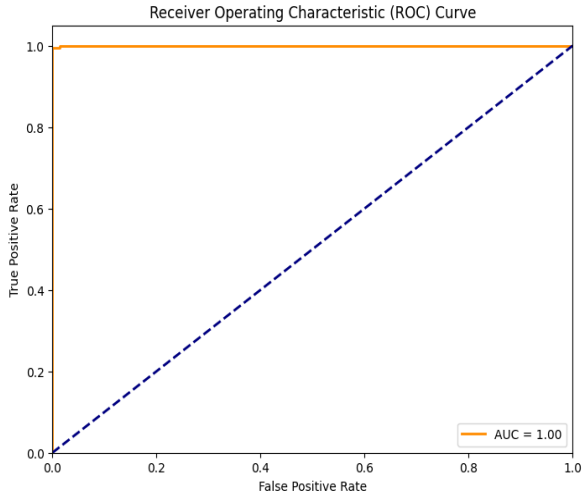


**Fig. 8.ROC Curve**

Visually, the receiver's operators (ROC) curve illustrates the rate of truly positive results (TPR) versus the rate of actual positives (FPR) at different threshold levels, therefore displaying the outcomes of the binary information classification. Lines in orange show the actual performance of a model; lines with crosswise dashed focusses of gravitation for at selected show Based on an Area As Figure 8 shows, in any "curve" (AUC) of 1.00 the model appropriately and error-free differentiates between the two groups.

**Conclusion and future work**

In conclusion WDBC is a useful dataset for breast cancer classification when employing a deep neural network. Our model displays great generalization and accurate differentiation with training 98.63% and Testing accuracy of 97.37%. The model is clearly successful based on strong recall, accuracy, and F1 score of 0.99. These findings imply the prospective importance of deep learning overall in allowing early identification-based better medical care for malignant tumor patients.

To improve the performance in the model even further, future research should investigate

complicated structures and examine them using more varied datasets. In practical clinical settings, this strategy will serve to enhance patient care, therefore benefitting physicians with respect to the diagnosis and prognosis for breast cancer.

## REFERENCES

Zhu, C. *et al.* Breast cancer histopathology image classification through assembling multiple compact CNNs. *BMC Med. Inform. Decis. Mak.* **19**(1), 1–17. https://doi.org/10.1186/s12911-019-0913-x (2019).

M. W. Arshad, "prediction and diagnosis of breast cancer using machine learning and ensemble classifiers," central asian jour-nal of mathematical theory and computer sciences, vol. 4, pp. 49-56, 2023.

Sengar, P. P., Gaikwad, M. J., & Nagdive, A. S. (2020, August). Comparative study of machine learning algorithms for breast cancer prediction. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 796-801). IEEE.

Wang, R. (2022, December). Comparison of decision tree, random forest and linear discriminant analysis models in breast cancer prediction. In Journal of Physics: Conference Series (Vol. 2386, No. 1, p. 012043). IOP Publishing.

Mangukiya, M., Vaghani, A., & Savani, M. (2022). Breast cancer detection with machine learning. International Journal for Research in Applied Science and Engineering Technology, 10(2), 141-145.

Mangukiya, M., Vaghani, A., & Savani, M. (2022). Breast cancer detection with machine learning. International Journal for Research in Applied Science and Engineering Technology, 10(2), 141-145.

Kadhim, R. R., & Kamil, M. Y. (2022). Comparison of breast cancer classification models on Wisconsin dataset. Int. J. Reconfigurable Embed. Syst. ISSN, 2089, 4864.

Monirujjaman Khan, M., Islam, S., Sarkar, S., Ayaz, F. I., Kabir, M. M., Tazin, T., ... & Almalki, F. A. (2022). [Retracted] Machine Learning Based Comparative Analysis for Breast Cancer Prediction. Journal of Healthcare Engineering, 2022(1), 4365855.

Sengar, P. P., Gaikwad, M. J., & Nagdive, A. S. (2020, August). Comparative study of machine learning algorithms for breast cancer prediction. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 796-801). IEEE.

Mehta, D., Mohite, A., Shinde, V., Khatri, R., & Dokare, I. (2022). Detection of Breast Cancer using Machine Learning Algorithms. Available at SSRN 4108758.

Singh, D., Nigam, R., Mittal, R., & Nunia, M. (2023). Information retrieval using machine learning from breast cancer diagnosis. Multimedia Tools and Applications, 82(6), 8581-8602.

Serajee, N. H., Mannan, S., Abdulghafor, R., Wani, S., Ibrahim, A. A., & Olowolayemo, A. (2022). Breast Cancer Prediction Using Machine Learning. International Journal on Perceptive and Cognitive Computing, 8(1), 24-28.

Jasim, A. A., Jalal, A. A., Abdulateef, N. M., & Talib, N. A. (2022). Effectiveness evaluation of machine learning algorithms for breast cancer prediction. Bulletin of Electrical Engineering and Informatics, 11(3), 1516-1525.

Moldovanu, S., Nastase, I. N. A., Miron, M., & Moraru, L. (2022). Performance comparison of two non-parametric classifiers for classification using geometric features. Analele Universităţii" Dunărea de Jos" din Galaţi. Fascicula II, Matematică, fizică, mecanică teoretică/Annals of the" Dunarea de Jos" University of Galati. Fascicle II, Mathematics, Physics, Theoretical Mechanics, 45(2), 59-62.

Ed-daoudy, A., & Maalmi, K. (2020). Breast cancer classification with reduced feature set using association rules and support vector machine. Network Modeling Analysis in Health Informatics and Bioinformatics, 9(1), 34.

Wang, X., Ahmad, I., Javeed, D., Zaidi, S. A., Alotaibi, F. M., Ghoneim, M. E., ... & Eldin, E. T. (2022). Intelligent hybrid deep learning model for breast cancer detection. Electronics, 11(17), 2767.

Ara, S., Das, A., & Dey, A. (2021, April). Malignant and benign breast cancer classification using machine learning algorithms. In 2021 International Conference on Artificial Intelligence (ICAI) (pp. 97-101). IEEE.

https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset/data

https://john.sisler.info/resume/deep-learning-specialization/neural-networks-and-deep-learning

https://market.us/report/breast-cancer-market/

American Cancer Society. *Breast Cancer Facts & Figures 2023–2024*. Atlanta: American Cancer Society, 2024.

World Health Organization (WHO). *Cancer: Key Facts*. Geneva: WHO, 2023.

National Cancer Institute (NCI). *Types of Tumors: Benign and Malignant*. Bethesda, MD: NCI, 2023.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. "Global Cancer Statistics 2023: GLOBOCAN Estimates of Incidence and Mortality Worldwide," *CA: A Cancer Journal for Clinicians*, vol. 73, no. 1, 2023, pp. 15–40.

Waks, A. G., & Winer, E. P. "Breast Cancer Treatment: A Review," *JAMA*, vol. 321, no. 3, 2023, pp. 288–300.