Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 18s (2025)
https://www.jneonatalsurg.com

# Prediction of Breast Cancer using Deep Learning Algorithms and Gradient Boosting

**Km Neha, Ashish Gupta, Govind Singh Panwar, Gurunath S Waghale, Parul Goyal**

[1]Department of Computer Science and Engineering, DIT University, Dehradun, India, nehapanwar2016@gmail.com
[2]Department of Computer Science and Engineering, Tula's Institute Dehradun, India, ashishgupta.wit@gmail.com
[3]Department of Computer Science and Engineering, Dev Bhoomi Uttarakhand University, Dehradun, India, govindpanwar@gmail.com
[4]Department of Computer Science and Engineering, School of Computing, MIT Art, Design and Technology University, Pune, India, guruwaghale111@gmail.com
[5]Department of Computer Science & Engineering, M. M. Engineering College, Maharishi Markandeshwar Deemed to be University, Mullana, Ambala – 133207, Haryana, India, parul.goyal@mmumullana.org

Cite this paper as: Km Neha, Ashish Gupta, Govind Singh Panwar, Gurunath S Waghale, Parul Goyal, (2025) Prediction of Breast Cancer using Deep Learning Algorithms and Gradient Boosting. *Journal of Neonatal Surgery*, 14 (18s), 380-387.

**ABSTRACT**

One of the main causes of death for women, particularly in poorer nations, is breast cancer. In order to lower death rates, timely diagnosis, detection, prediction, and effective treatment are now essential. Models for predicting and diagnosing breast cancer become more reliable and accurate as artificial intelligence, machine learning, and deep learning approaches are used more frequently. Examining the efficacy of various machine learning and contemporary deep learning models for breast cancer diagnosis and prediction is the goal of this study. This study contrasts cutting-edge approaches that make use of deep learning models with conventional machine learning classification methods. Deep learning models like Neural Decision Forest and Multilayer Perceptron were employed, along with well-known classification models like k-Nearest Neighbors (kNN), Gradient Boosting, Support Vector Machine (SVM), Neural Network, CN2 rule inducer, Naive Bayes, Stochastic Gradient Descent (SGD), and Tree. The Orange and Python tools were used to conduct the experiment, which assesses their diagnostic efficacy in detecting breast cancer. Transparency and accessibility in the study strategy are made possible by the evaluation's usage of UCI's publicly available Wisconsin Diagnostic Data Set. Result: In both malignant and benign instances, the mean radius ranges from 6.981 to 28.110, and the mean texture ranges from 9.71 to 39.28. SVM has the lowest accuracy and sensitivity at 88%, whereas gradient boosting and CN2 rule inducer classifiers do better. With an AUC value of 0.98%, the CN2 rule inducer classifier obtains the highest ROC curve score for both benign and malignant breast cancer datasets. With a higher AUC-ROC of 0.9959, accuracy of 96.49%, precision of 96.57%, recall of 96.49%, and F1-Score of 96.50%, MLP displays can differentiate between positive and negative classes. GB and the CN2 rule outperformed the other models among the most popular classifier models. Deep learning's MLP, however, yielded the best overall results.

**Keywords:** Deep Learning, AI, Breast Cancer XAI, Transfer learning, Mammography.

## 1. INTRODUCTION

Every nation in the world faces a major obstacle to raising life expectancy and a major contributor to mortality: cancer. Female breast cancer demonstrated a considerable increase with 2.3 million new cases, showing a parallel rise in death rates. Women's lives are uniquely impacted by breast cancer [1]. Mortality may be decreased by employing tactics including proactive diagnosis, early detection, and more awareness [2]. In today's medical environment, collecting large amounts of data on a variety of illnesses improves research and enables predictive insights. Technology helps pathologists and doctors make precise forecasts, prevent wasteful medical expenses, and guarantee the best possible care.

Sometimes, early detection is a life-saving measure [3]. Breast cancer diagnosis and prediction depend on a variety of machine learning classification algorithms. Numerous studies highlight the necessity of employing diverse strategies to address challenges in breast cancer prediction. Several data mining machine learning algorithms using the Wisconsin Breast Cancer diagnostic dataset can help pathologists better recognize trends in disease progression [4]. Significant progress has been made in the analysis of health data using machine learning algorithms, particularly in relation to breast cancer. To predict cancer susceptibility, recurrence, survivorship, and treatment outcomes, algorithms like random forest, extreme learning machine, naive Bayes, artificial neural networks, and support vector machine have been applied to the Breast Cancer Wisconsin Data Set [5]. Significant progress has been achieved in the identification and diagnosis of breast cancer, which has improved patient outcomes by enabling timely intervention and individualized treatment plans. In the early phases, machine learning algorithms have demonstrated potential in identifying the aggressiveness of breast cancers and forecasting survival [6]. This study offers numerous benefits since it supports prompt diagnosis and treatment approach change. Additionally, by influencing factors like future research, the affordability, and worldwide accessibility of healthcare services, it has the potential to change the current state of breast cancer management. Reducing healthcare costs will be aided by the use of machine learning techniques for breast cancer screening. However, it also has the potential to significantly and favorably influence breast cancer treatment globally.

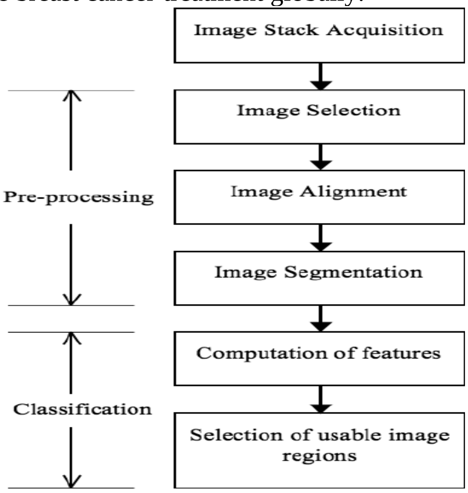



**Fig 1: Basic Preprocessing of Images**

The occurrence of breast cancer maintains its position as the main malignant disease that impacts women worldwide substantially. The early discovery of breast cancer remains vital for enhancing both treatments and patient survival results. AI technology together with deep learning methods has transformed medical imaging because it enables accurate automated analysis systems. The application of CNNs has delivered outstanding performance when used to identify breast cancer in mammography and histopathology images [5]. Modern imaging systems combined with artificial intelligence have not eliminated the persistent problems with secure and precise breast cancer diagnosis.

The combination of effectiveness with deep learning models remains limited by their black box nature because their complex decision-making and architectural elements remain unclear. Clinical decision support requires explainable systems due to which uninterpretable systems face hurdles for trust-based adoption in medical settings [6]. Medical imaging datasets contain different types of data variability across quality levels and both patient-related factors and pathological information. The inconsistent characteristics of medical images create obstacles for creating AI models which can perform reliably and accurately on various patient communities [7]. The goal of Explainable AI methods is to improve both transparency and interpretability features of AI systems particularly within medical applications. XAI techniques supply clinicians with clear model explanation capabilities and feature analysis which helps them understand and build trust in AI diagnostic systems. The understanding serves to enable AI's integration into medical operations and supports AI system collaboration with medical staff professionals [8]. Transfer learning has become a vital method which helps remedy the difficulties related to scant annotated medical imaging resources. Detected datasets of ImageNet allow transfer learning techniques to update neural networks for processing smaller domain-specific image collections like mammography images [9]. The network of DRCNNs combines both CNN-based spatial hierarchical knowledge with RNN-based sequential connection detection [10]. The hybrid model architecture demonstrates excellent suitability to process sequential medical data while providing satisfactory results for breast cancer detection tasks from longitudinal imaging examinations [11]. This study aims to Explainable AI techniques in conjunction with transfer learning and DRCNN:

- Rephrase the detection performance of breast cancer in mammography imaging using new methods.
- The AI system's decision process should be explained through visual representations to make its operations more understandable.
- Tests of the proposed method must be performed on various datasets to confirm operational performance across multiple imaging environments as well as demographic populations.

The main body of the research paper follows this introduction section along with its structure and part 2 contains the literature review analysis. We explain our research methods at the high-level in the third section of this document. Section IV provides detailed explanations about implementation approach and performance measures while Section V shows results and discussions for the current research.

## 2. LITERATURE SURVEY

In the medical field, machine learning techniques have been applied in a number of recent studies to diagnose breast cancer. Other academics have modified these algorithms to handle complex scenarios since they produce satisfying results. For the prediction and diagnosis of carcinoma breast, ongoing research offers insightful recommendations. kNN, NB, Tree, SVM, and logistic regression are among the machine learning techniques for breast cancer diagnosis that have been the subject of numerous research papers. According to a study, early identification is crucial for enhancing the prognosis and survival of breast cancer. In order to create models that can accurately identify breast cancer as either benign or malignant, they placed a strong emphasis on the use of classification algorithms [7]. Another study demonstrated positive results using upgraded machine learning algorithms, which resulted in better breast cancer treatment options and increased proficiency in less invasive predictive medicine [8]. In addition, a study that predicted the likelihood of a cancerous breast used the RF and XG Boost models, which had accuracy rates of 74.73% and 0.73, respectively [9]. In order to improve survival rates and forecast the development of breast cancer, another study employed the EXSA GB technique [10]. In one study, an ensemble model comprising kNN, SVM, and DT was created, and it achieved 78% accuracy [11]. Additionally, a different study employed a number of supervised learning algorithms on the Wisconsin dataset, with the Artificial Neural Network (ANN) achieving the highest accuracy of 98.57% [5]. Additionally, the highest accuracy was obtained by utilizing SVM and a quadratic kernel [12]. Depending on the particulars of the correlations between dataset features, either linear or nonlinear approaches can be used to reduce data [13]. Support Vector Machines (SVM), logistic regression, naive Bayes, and random forest were evaluated to find commonalities and differences [14]. Another study used the Wisconsin breast cancer dataset as a reference. RF produced the best results with a precision of 0.997 and the least amount of error when Anaconda Data Science boards were utilized [15, 16]. A neural network (NN) approach is used to categorize different types of breast cancer, with an emphasis on the MLP. The neural network's main job is to classify the input data into two categories: benign and aggressive breast cancer. A method of collective erudition, known for its efficacy, is used to strike a balance between bias and change. Combining different classifiers to create a single classification model improves classification performance; this idea has been proven in numerous experimental studies.

Three basic methods are used in ensemble classification: bagging, boosting, and stacking. For instance, stacking entails integrating the output of numerous categorization models into a single model [17]. It found a dearth of research assessing how well deep learning and machine learning algorithms identify breast cancer. Researchers got interested in contrasting the effectiveness of conventional and contemporary breast cancer prediction models after realizing this disparity.

## 3. CONFIGURATION AND TECHNIQUES

The main idea of the proposed breast cancer detection system is shown in Figure 1. Despite the fact that many classification approaches are used to breast cancer data, each classifier performs differently on the same dataset. Consequently, boosting and bagging are employed as part of an ensemble technique. This approach learns from prior classifiers and combines data from several classifiers. Gathering information is the first step in doing this. The data is then preprocessed in order to pick attributes. 20% of the dataset is used for testing, and the remaining 80% is used for training. Labeled data with benign and malignant classifications are included in the collection. A model is then created utilizing the training data and a number of supervised classification techniques. Several classifiers are used to examine the test data, and their respective performances are compared with deep learning models and to each other.

This study made use of the Wisconsin Malignant Breast Diagnostic dataset, which is publicly accessible via the UCI ML repository. Data from 569 people with malignant breast illnesses is part of the collection. In order to analyze data and detect breast cancer in this study, researchers employed a variety of machine learning algorithm models. A subset of AI called machine learning (ML) is used to classify data using models that have been developed, most notably in the analytics of breast cancer prediction. It offers automated methods for analyzing big datasets. Machine learning (ML) techniques were employed in this study, and a scientific dataset of patients with breast cancer was sourced from Kaggle (https://www.kaggle.com/uciml/breast-cancer-wisconsin-data). The data was assessed using a number of criteria from digitally available images of breast mass aspirations using fine needles. In particular, these characteristics were employed to distinguish between benign and malignant tumors.

In order to train and test the machine in an 80:20 ratio, the data were randomly separated into two sets using the orange tool. The accuracy of the diagnosis and the efficacy of medical therapy are important aspects that affect a patient's chances of surviving cancer and avoiding recurrence. After this division, the model is trained using the designated training sets, and its effectiveness is assessed using test data. A person's likelihood of being impacted is determined by a number of feature variables. Gathering necessary data for pre-processing to improve its quality is the first stage, and this is accomplished using a variety of pre-processing techniques. It involves simply selecting the necessary properties.

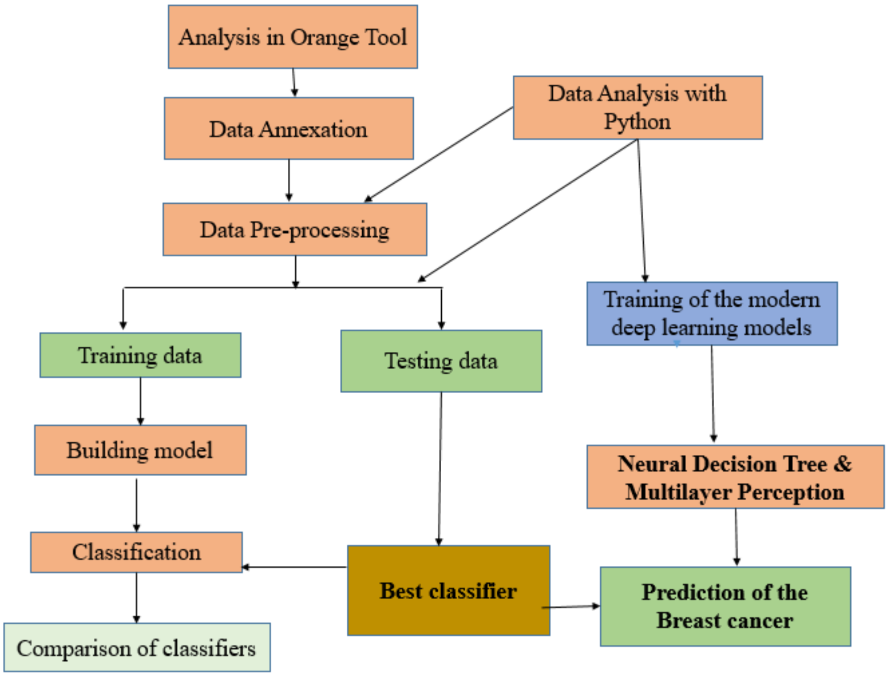



**Figure 2: Proposed Flow**

After the data is prepared, it is trained using a variety of machine models. Test data is used to assess the model's performance following training. This method uses an ensemble classifier along with a variety of machine learning algorithms. A number of machine models are used to train the data after it has been prepared. Test data is then used to evaluate the model's performance. This method uses a collective categorization model in conjunction with other machine learning methods. The goal of collective erudition is to combine the results of several learning algorithms to produce a better predictive model. It is accomplished by combining different supervised learners to increase the prediction power of the model. Several techniques, such as kNN, gradient boosting, SVM, neural networks, CN2 rule inducer, naive Bayes, tree, and SGD, are used in this situation. A pre-processed dataset of breast cancer is used to predict malignancy using a variety of classification and deep learning techniques.

Predictive models for breast cancer

kNN: This computes distances for label assignment and uses all training data for similarity-based classification.

Hyperplanes without the necessary information for the previous distribution are used in SVM.

NB: This approach makes predictions based on likelihoods and is based on the Bayes theorem.

DT: This method divides occurrences using the Gini Index or Information Gain and classifies them according to feature values. Leaf nodes stand in for labels.

Gradient Boosting: This technique improves overall forecast accuracy by gradually building an ensemble of weak learners, usually decision trees, and reduces errors by highlighting cases with previous model defects.

Neural Networks: This model, which is made up of connected nodes or neurons and was inspired by the human brain, is excellent at learning intricate patterns and correlations, which makes it helpful for tasks like regression and classification.

Stochastic Gradient Descent: SGD improves training efficiency and convergence by iteratively adjusting model parameters to minimize the loss function.

Multilayer Perceptrons (MLP) are multilayer neural networks that can accomplish challenging tasks by using nonlinear activation functions.

Ensembles of decision trees or small neural networks called neural decision forests are intended to improve interpretability, performance, and nonlinearity. To enhance model capabilities, they integrate neural network principles with components from traditional decision forests.

The best machine-learning classification techniques discussed in this research are found using the Orange mining tool. This platform facilitates the development of several algorithms, enabling quick dataset exploration and analysis as well as contributing to the computational framework and analytical results of the study. The effectiveness of deep learning models was assessed by researchers using the Python software framework. The accuracy and Receiver Operating Characteristic Area Under the Curve (ROC AUC) measurements were carefully examined in a comprehensive comparison.

## 4. RESULT AND DISCUSSION

This study examined data from 569 patients with breast cancer diagnoses from the Wisconsin Hospital. Out of these patients, thirty-seven percent had malignant cases diagnosed, while sixty-three percent had benign conditions. the properties of cell nuclei that could be important when diagnosing breast cancer. The characteristics that characterize the typical size and shape of cell nuclei are Radius_Mean, Texture_Mean, Perimeter_Mean, and Area_Mean. The Worst Features—Radiaus_Worst, Perimeter_Worst, and Area_Worst—represent the most undesired (largest) size and form characteristics of cell nuclei and are shown in the table. For instance, mean texture ranges from 9.71 to 39.28, while mean radius ranges from 6.981 to 28.110, with a mean value of 14.12729. As indicated in Table 2, all hyper-parameters were adjusted and the prediction was executed; the CN2 introducer and Gradient Boosting produced superior prediction output compared to the other predictors (Table 1). Metrics like accuracy, AUC, precision, and recall were used to evaluate the performance of several popular machine learning classification and deep learning models that were applied to the Wisconsin Breast Cancer dataset. The model's performance was specifically assessed by the researcher using the AUC statistic. The results were separated into two categories: ensemble and conventional machine learning techniques.

**Table I. Distinct Geometrical and Textural Features**

| Features | Mean | Dispersion | Minimum Value | Maximum Value |
|---|---|---|---|---|
| Radius_Mean | 14.12729 | 0.24923 | 6.981 | 28.11 |
| Texture_Mean | 19.2896 | 0.2228 | 9.71 | 39.28 |
| Perimeter_Mean | 91.969 | 0.264 | 43.79 | 188.5 |
| Area_Mean | 654.889 | 0.537 | 143.5 | 2501 |
| Concavity_Mean | 0.0887993 | 0.896963 | 0 | 0.4268 |
| Concave Points_Mean | 0.0489191 | 0.792506 | 0 | 0.2012 |
| Symmetry_Mean | 0.181162 | 0.151192 | 0.106 | 0.304 |
| Fractal_Dimension_Mean | 0.0627976 | 0.1123316 | 0.04996 | 0.09744 |
| Radius_Worst | 16.26919 | 0.29682 | 7.93 | 36.04 |
| Perimeter_Worst | 107.2612 | 0.646 | 185.2 | 4254 |
| Area_Worst | 880.583 | 0.646 | 185.2 | 4254 |
| Smoothness_Worst | 0.1323686 | 0.1723396 | 0.07117 | 0.2226 |
| Symmetry_Worst | 0.290076 | 0.213093 | 0.1565 | 0.6638 |
| Diagnosis | Benign=65% | | | Malignancy =35% |

**Table II. Performance evaluation for classification models**

| Model | AUC | CA | F1 | Precision | Recall | MCC |
|---|---|---|---|---|---|---|
| kNN | 0.992 | 0.907 | 0.903 | 0.919 | 0.907 | 0.808 |
| Gradient Boosting | 1 | 1 | 1 | 1 | 1 | 1 |
| SVM | 0.992 | 0.889 | 0.884 | 0.906 | 0.889 | 0.773 |
| Neural Network | 0.988 | 0.895 | 0.89 | 0.91 | 0.895 | 0.783 |
| CN2 rule inducer | 1 | 1 | 1 | 1 | 1 | 1 |
| Naive Bayes | 0.985 | 0.942 | 0.942 | 0.942 | 0.942 | 0.876 |
| Tree | 0.988 | 0.989 | 0.989 | 0.99 | 0.989 | 0.977 |
| SGD | 0.985 | 0.989 | 0.988 | 0.988 | 0.988 | 0.974 |

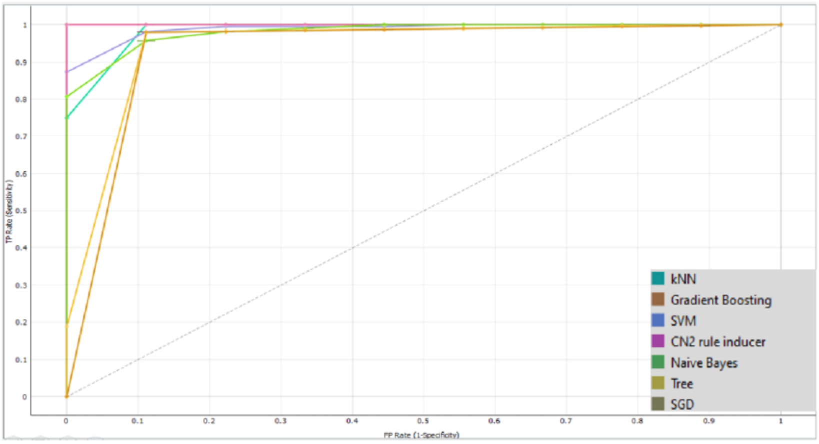



**Fig. 3. ROC Analysis of Malignant Tumour**

**Table III. Prediction Output Using Deep Leaning Model (NDF and MLP)**

| Metrics | Neural Decision Forest | Multilayer Perceptron |
|---|---|---|
| AUC-ROC | 0.9667 | 0.9959 |
| Accuracy | 95.61% | 96.49% |
| Precision | 100% | 96.57% |
| Recall | 89.36% | 96.49% |
| F1-Score | 94.38% | 96.50% |

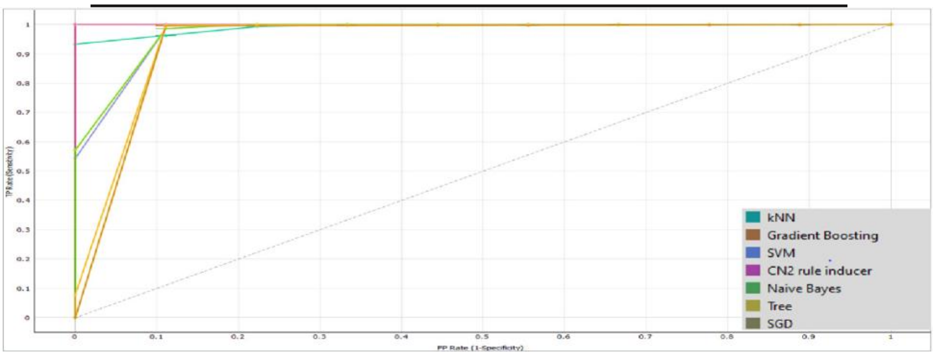



**Fig. 4. ROC Analysis of benign Tumour**

Predicted

| Actual | B | M | Σ |
|---|---|---|---|
| B | 87.1 % | 0.0 % | 357 |
| M | 12.9 % | 100.0 % | 212 |
| Σ | 410 | 159 | 569 |

Predicted

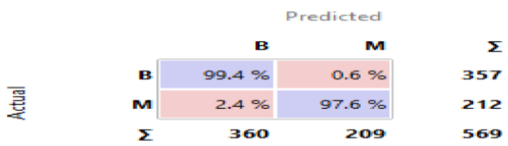
| Actual | B | M | Σ |
|---|---|---|---|
| B | 100.0 % | 0.0 % | 357 |
| M | 0.0 % | 100.0 % | 212 |
| Σ | 357 | 212 | 569 |

Predicted

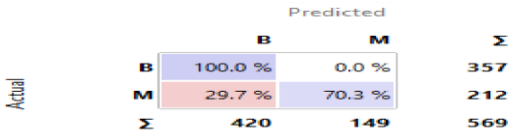
| Actual | B | M | Σ |
|---|---|---|---|
| B | 95.2 % | 4.8 % | 357 |
| M | 7.5 % | 92.5 % | 212 |
| Σ | 356 | 213 | 569 |

Predicted

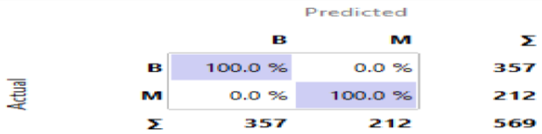
| Actual | B | M | Σ |
|---|---|---|---|
| B | 99.7 % | 0.3 % | 357 |
| M | 2.4 % | 97.6 % | 212 |
| Σ | 361 | 208 | 569 |



| Actual | Predicted B | Predicted M | Σ |
|---|---|---|---|
| B | 99.4 % | 0.6 % | 357 |
| M | 2.4 % | 97.6 % | 212 |
| Σ | 360 | 209 | 569 |



| Actual | Predicted B | Predicted M | Σ |
|---|---|---|---|
| B | 100.0 % | 0.0 % | 357 |
| M | 29.7 % | 70.3 % | 212 |
| Σ | 420 | 149 | 569 |



| Actual | Predicted B | Predicted M | Σ |
|---|---|---|---|
| B | 100.0 % | 0.0 % | 357 |
| M | 0.0 % | 100.0 % | 212 |
| Σ | 357 | 212 | 569 |

**Fig. 5. Confusion matrix for kNN, Gradiant Boosting, SVM, CN2 Rule inducer, Naive Bayes, Tree, SGD**

## 5. CONCLUSION

This study examines two deep-learning models and seven distinct classifications for breast cancer categorization using the Breast Cancer Wisconsin (diagnostic) dataset. The Orange data mining tool and Python program are utilized for feature selection, while the Standard Scaler module is used for data preparation. The deep learning models were constructed using NDF and MLP approaches, while the machine learning models were constructed using kNN, Gradient Boosting, SVM, Neural Network, CN2 rule introducer, Naive Bayes, Decision Tree, and SGD. The evaluation measures performance markers such accuracy, AUC, precision-recall curve, sensitivity, and f1-score and associates expected and actual outcomes using a confusion matrix. After processing the data in the tool, this study found a drop in the maximum area mean and smoothness worst values, suggesting a potential rise in false positives. To comprehend how different features relate to a patient's prognosis, it is essential to look at the relationships between various components of cancer of the beast detection. After naive Bayes, DT, SGD, and kNN models, gradient boosting and CN2 rule introducer models consistently show the highest efficiency. Breast cancer is a prevalent condition that impacts women worldwide. It has the potential for machine-learning algorithms to have a transformative impact on early detection and prognosis. Invasive and ductal in situ are the two types of ductal carcinoma. Effective therapy depends on timely detection, highlighting the importance of using accurate screening techniques. USG and mammography are two common radiographic techniques used to identify breast cancers early on. Deep learning algorithms that can detect breast cancer in digital mammography have been made possible by significant advancements in artificial intelligence, which have significantly increased mammography precision. To further enhance image processing and lessen the need for human visual identification in the detection of breast cancer, an MRI of the breast serves as an imaging intelligence integrated tool that has emerged in the field of patient care. Future machine learning studies in breast cancer diagnosis could look into a variety of topics, such as improving on current models and implementing novel techniques. The ongoing cooperation of data scientists, medical professionals, and researchers is essential to the ongoing advancement of breast cancer detection and therapy. A fuller understanding of the illness is aided by ongoing efforts to increase the accuracy and efficiency of diagnostic tools. There is increasing potential for innovative applications to improve patient outcomes as technology develops [18]. Advances in breast cancer diagnosis and therapy continue to depend heavily on the interplay between scientific proficiency and technical innovation. Finally, the Breast Cancer Wisconsin dataset was utilized to evaluate a number of deep learning and machine learning models for the categorization of breast cancer. Evaluation metrics showed how effective models like the CN2 rule introducer and gradient boosting were. The study underscored the significance of investigating correlations in the screening process for breast cancer, specifically the possible impact of machine-learning algorithms on early diagnosis. Current developments in artificial intelligence, particularly in the fields of breast MRI and mammography, point to exciting opportunities to improve diagnostic precision. In order to enhance breast cancer detection and treatment, future research should focus on enhancing current models and promoting cooperation between data scientists and medical specialists.

Data: The data presented in this research may be found in the UCI Machine Learning Repository, which can be found at https://archive.ics.uci.edu/dataset/17/breast+cancer+w isconsin+diagnostic.

## REFERENCES


[1] Bayrak EA, Kırcı P, Ensari T, editors. Comparison of machine learning methods for breast cancer diagnosis. 2019 Scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT); 2019: Ieee. https://doi. org/10.1109/EBBT.2019.8741990

[2] Patil S, Kirange D, Nemade V. Predictive modelling of brain tumor detection using deep learning. J. Crit. Rev. 2020;7(04):1805-13.

[3] Eroğlu Y, Yildirim M, Çinar A. Convolutional neural networks based classification of breast ultrasonography images by hybrid method with respect to benign, malignant, and normal using mrmr. Comput Biol Med.

2021;133:104407. https:// doi.org/10.1016/j.compbiomed.2021.104407.

[4] Dora L, Agrawal S, Panda R, Abraham A. Optimal breast cancer classification using gauss–newton representation based algorithm. Expert Syst Appl. 2017;85:134-45. https:// doi.org/10.1016/j.eswa.2017.05.035

[5] Islam MM, Haque MR, Iqbal H, Hasan MM, Hasan M, Kabir MN. Breast cancer prediction: A comparative study using machine learning techniques. SN Computer Science. 2020;1:1-14.

[6] Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. Radiology. 2018;286(3):800-9. https://doi.org/10.1148/ radiol.2017171920.

[7] Priya R VS. Bio-inspired ensemble feature selection (biefs) and kernel extreme learning machine classifier for breast cancer diagnosis. Int J Health Sci. 2022;6(S5):1404-29.

[8] Kim J, Shin H. Breast cancer survivability prediction using labeled, unlabeled, and pseudo-labeled patient data. J Am Med Inform Assoc. 2013;20(4):613-8. https://doi. org/10.1136/amiajnl-2012-001570.

[9] Kabiraj S, Raihan M, Alvi N, Afrin M, Akter L, Sohagi SA, Podder E, editors. Breast cancer risk prediction using xgboost and random forest algorithm. 2020 11th international conference on computing, communication and networking technologies (ICCCNT); 2020: IEEE. https:// doi.org/10.1109/ICCCNT49239.2020.9225451

[10] Liu P, Fu B, Yang SX, Deng L, Zhong X, Zheng H. Optimizing survival analysis of xgboost for ties to predict disease progression of breast cancer. IEEE Trans Biomed Eng. 2021;68(1):148-60. https://doi.org/10.1109/ tbme.2020.2993278.

[11] Nanglia S, Ahmad M, Khan FA, Jhanjhi N. An enhanced predictive heterogeneous ensemble model for breast cancer prediction. Biomedical Signal Processing and Control. 2022;72:103279.

[12] Obaid OI, Mohammed MA, Ghani MKA, Mostafa A, Taha F. Evaluating the performance of machine learning techniques in the classification of wisconsin breast cancer. International Journal of Engineering & Technology. 2018;7(4.36):160-6.

[13] Abunasser BS, AL-Hiealy MRJ, Zaqout IS, Abu-Naser SS. Breast cancer detection and classification using deep learning xception algorithm. Int J Adv Comput Sci Appl. 2022;13(7). https://doi.org/10.14569/IJACSA.2022.0130729

[14] Kharya S, Soni S. Weighted naive bayes classifier: A predictive model for breast cancer detection. Int J Comput Appl. 2016;133(9):32-7. https://doi.org/10.5120/ijca2016908023

[15] Huang Q, Chen Y, Liu L, Tao D, Li X. On combining biclustering mining and adaboost for breast tumor classification. IEEE Transactions on Knowledge and Data Engineering. 2019;32(4):728-38. https://doi.org/10.1109/ TKDE.2019.2891622

[16] Sivapriya J, Kumar A, Sai SS, Sriram S. Breast cancer prediction using machine learning. International Journal of Recent Technology and Engineering (IJRTE). 2019;8(4):4879-81.

[17] Tang X, Cai L, Meng Y, Gu C, Yang J, Yang J. A novel hybrid feature selection and ensemble learning framework for unbalanced cancer data diagnosis with transcriptome and functional proteomic. IEEE Access. 2021;9:51659-68. https://doi.org/ 10.1109/ACCESS.2021.3070428

[18] Patil S, Moafa IH, Alfaifi MM, Abdu AM, Jafer MA, Raju L, et al. Reviewing the role of artificial intelligence in cancer. Asian Pac J Cancer Biol 2020;5(4):189-99. 10.31557/ apjcb.202