ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

A Comparative Analysis And Classification Of Breast Cancer Using Machine Learning

P. Lahari [1], P. Sarath [2], R. Manoj Kumar [3], M. Kishore [4], D. Balaji [5], M. Divya [6],*
Department of Computer Science and Engineering [123456],
Aditya Institute of Technology and Management [123456],
Andhra Pradesh, Tekkali 532201.
divya.modalavalasa@gmail.com

Abstract
Breast cancer is an especially dangerous disease that threatens the health of women. It happens when breast cells become cancerous that can move to other areas within the body. The earlier breast cancer is recognized, the greater the odds of therapeutic success. The application of machine learning in the early diagnosis of breast cancer is the main objective of this literature review. Specifically, this paper compares five different machine learning algorithms, namely Gradient Boosting (GB), Extreme Gradient Boosting (XGB), Light Gradient Boosting (Light GBM), Decision Tree (DT), and K-Nearest Neighbors (KNN) on the Wisconsin Diagnostic Breast Cancer (WDBC) dataset. The dataset is divided into two sections for training and testing the machine learning algorithms. After careful consideration of the best algorithm, the model then classify cancer as either benign or malignant.

Keywords: Breast Cancer, Gradient Boosting, Extreme Gradient Boosting, Light Gradient Boosting, Decision Tree, K-Nearest Neighbors, Benign, Malignant

1. Introduction

Breast cancer is a prevalent type of cancer among women. Accurate methods have the potential to raise survival rates to more than 86%. Early identification of breast cancer is essential for enhancing patient survival rates. The development of cancerous malignant lumps from breast cells characterizes breast cancer. Unfortunately, doctors may misdiagnose benign tumours, which are noncancerous, as malignant. To address this issue, computer-aided detection (CAD) systems that utilize machine learning are needed to provide accurate breast cancer diagnosis. CAD systems can aid in the early detection of breast cancer, which is critical in increasing the chances of survival by providing better treatment. Improving existing approaches to predict breast cancer at an early stage is necessary to reduce the high death rate associated with breast cancer.

There are two type of breast cancer as follows:

1. Benign Tumours
2. Malignant Tumours

Benign Tumours: Non-Cancerous

A benign tumour is one which does not include malignant cells and lacks the tendency to grow

Cite this article as: P.Lahari, P.Sarath, R.Manoj Kumar, M.Kishore, D.Balaji & M.Divya, " A Comparative Analysis And Classification Of Breast Cancer Using Machine Learning", International Journal & Magazine of Engineering, Technology, Management and Research (IJMETMR), ISSN 2348-4845, Volume 10, Issue 4, April 2023, Page 27-39.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

or infect surrounding tissues. Except when they cause damage to adjacent tissues, nerves, or blood vessels, benign tumours are relatively less threatening. Uterine fibroids and lymphomas are examples of benign tumours. Surgery might be required in some instances to remove benign tumours, which can grow to be quite gigantic, weighing several pounds. Benign tumours could eventually push on or restrict essential organs. Specific benign tumours, such as intestinal polyps, are designated precancerous and are treated with surgery to stop cancerous development. Whereas the benign tumours often reappear after removal, when they do, they ordinarily reappear in the same spot.

Malignant Tumours: Cancerous

Malignant tumors are made up of cancer cells that have the ability to invade and penetrate surrounding tissues. These cancer cells can spread to other areas of the body by traveling through the bloodstream or lymph nodes. The disease can develop in different parts of the body, including the breast, lungs, intestines, reproductive system, skin, and blood. If breast cancer spreads to the lymph nodes, the cancer cells can travel to other organs, such as the liver or bones, and form more tumors. Biopsies of these tumors often show similarities to the primary breast cancer tumor. Therefore, early detection and treatment of malignant tumors is essential to prevent metastasis and improve patient outcomes.

Signs and Symptoms of Breast Cancer

Breast swelling

Discharge from the nipples

Dimpling of the skin

Nipple that turns inward

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Swelling or a lump under the collarbone

Skin changes on the breast or nipple

Fig-1: Signs and Symptoms of Breast Cancer

3. Literature Survey

In a recent study, researchers Sharma, Aggarwal, and Choudhury [1] have utilized machine learning algorithms to detect breast cancer. By examining different machine learning models and their ability to identify breast cancer from medical imaging data, the team evaluated model accuracy using sensitivity, specificity, and accuracy metrics while comparing their findings to traditional methods of breast cancer diagnosis. The team's results indicate that machine learning algorithms can be useful in detecting breast cancer and may even help radiologists make more precise diagnoses.

Mohammed, Darrab, et al. [2] delve into the various stages involved in detecting breast cancer, including image pre-processing, feature extraction, and classification. The team also compares the effectiveness of different machine learning algorithms, such as Artificial Neural Networks, Random Forest, and Support Vector Machines, in identifying breast cancer from medical imaging data. To evaluate the accuracy of the models, the authors use sensitivity, specificity, and F1 score, and compare their findings to traditional methods of breast cancer diagnosis. The study concludes that machine learning algorithms can offer precise detection of breast cancer and may even aid radiologists in making more accurate diagnoses. Nonetheless, the paper highlights the limitations of current machine learning techniques and suggests future research areas for improvement. Overall, this study provides a comprehensive overview of the current state of research on breast cancer detection using machine learning techniques.

Tahmooresi, et al. [3] conducted a study to explore the effectiveness of machine learning algorithms in the early detection of breast cancer. Published in the Journal of Telecommunication, Electronic and Computer Engineering (JTEC), the authors reviewed past studies related to breast cancer detection and machine learning. They discussed the challenges associated with early detection and emphasized the importance of precise and timely diagnosis for optimal patient outcomes. The authors then presented their own study, which utilized three different machine learning algorithms (k-NN, SVM, and Random Forest) to classify breast cancer cases as malignant or benign. The dataset consisted of 569 instances with 30 features extracted from breast tissue images. The results indicated that all three algorithms achieved high accuracy rates, with

Random Forest performing the best at 96.49%. The authors concluded that machine learning algorithms have potential for early detection of breast cancer and suggest further research to explore their application in clinical settings.

Khourdifi and Bahaj, et al. [4] conducted a study to evaluate the effectiveness of various machine learning algorithms for predicting and classifying breast cancer cases. Presented at the 2018 International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), the authors conducted a literature review on previous studies related to breast cancer diagnosis and machine learning algorithms. They emphasized the importance of early detection and precise diagnosis of breast cancer for improving patient outcomes. The authors then presented their own study, which involved using five different machine learning algorithms (SVM, Naive Bayes, Decision Tree, Random Forest, and K-NN) to predict and classify breast cancer cases as malignant or benign. The dataset used in the study contained 569 instances with 30 features extracted from breast tissue images. Results revealed that SVM and Random Forest achieved the highest accuracy rates, with SVM outperforming all other algorithms with an accuracy of 98.95%. The authors concluded that machine learning algorithms hold potential for breast cancer prediction and classification, and further research is necessary to explore their application in clinical settings.

Ebru Aydındag Bayrak [5] and colleagues from the Department of Computer Engineering at Istanbul University in Turkey reviewed two commonly used machine learning techniques for Wisconsin Breast Cancer classification. After evaluating the performance metrics of the applied machine learning techniques, the researchers found that Support Vector Machine Algorithm (SVM) demonstrated the highest accuracy of 96% for the diagnosis and prediction of the WBC dataset.

Nallamla, et al. [6] conducted a literature survey on the application of data mining techniques in cancer treatment. The study was published in the International Journal of Engineering and Technology (UAE). The authors conducted an extensive literature review on previous studies related to cancer treatment and data mining. They discussed the challenges associated with cancer treatment and the significance of effective treatment for improved patient outcomes. The authors also emphasized the potential of data mining techniques to aid in cancer treatment. The authors identified various data mining techniques that have been utilized in cancer treatment, including clustering, classification, association rule mining, and regression analysis. They discussed the advantages and limitations of each technique and provided examples of studies that have used these techniques to enhance cancer treatment. The authors concluded that data mining techniques can be effective in cancer treatment and suggest further research to explore their potential in clinical settings. They also highlighted the need for collaboration between clinicians and data scientists to develop effective data mining models for cancer treatment.

Alarabeyyat and Alhanahnah [7] conducted a study to investigate the effectiveness of the K-

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Nearest Neighbor (KNN) machine learning algorithm in detecting breast cancer. Presented at the 2016 International Conference on Developments in eSystems Engineering (DeSE), the authors conducted a literature review on previous studies related to breast cancer detection and machine learning algorithms. They highlighted the challenges associated with early detection of breast cancer and the significance of accurate diagnosis for improving patient outcomes. The authors emphasized the potential of machine learning methods to aid in breast cancer diagnosis. The authors then presented their own study, which utilized the k-NN algorithm to classify breast cancer cases as malignant or benign. The dataset used in the study contained 569 instances with 30 features extracted from breast tissue images. The results indicated that the KNN algorithm achieved high accuracy rates, with an accuracy of 94.74%. The authors concluded that the KNN algorithm could be effective in detecting breast cancer, and further research is necessary to assess its feasibility in clinical practice.

Mangukiya, M., Vaghani, et al. [8] conducted a study on machine learning techniques to facilitate the early detection of breast cancer. They performed data visualization and compared the performance of different machine learning algorithms, including Support Vector Machine (SVM), Decision Tree, Naive Bayes (NB), K Nearest Neighbours (k-NN), Adaboost, XGboost, and Random Forest, on the Wisconsin Breast Cancer Dataset. The study aimed to assess the accuracy of the classification of data in terms of the efficiency and effectiveness of each algorithm, measured through accuracy, precision, sensitivity, and specificity. The authors' objective was to review various techniques for detecting breast cancer early, efficiently, and accurately using machine learning. Experimental results revealed that XGboost exhibited the highest accuracy (98.24%) with the lowest error rate.

Ayush Sharma, et al. [9] conducted research on Machine Learning Approaches for Breast Cancer Diagnosis and Prognosis with the aim of predicting breast cancer as benign or malignant using a dataset from the Wisconsin Breast Cancer Data. The study employed sophisticated classifiers such as Logistic Regression, K-Nearest Neighbor, and Support Vector Machines to calculate the probability of recurrence in affected patients. The relationship between precision, recall, and the number of features in the dataset was also determined. The research involved clinical examination to detect the tumour/lump in the breast using imaging techniques such as Mammography, followed by FNA on the detected lump. The FNA procedure provided attributes such as tumour size, radius, area, etc., as features for modelling the data into a classifier. The study trained models such as SVM, K-nearest neighbor, and Logistic Regression using the data to make predictions.

The issue of detecting breast cancer from a set of symptoms has attracted the attention of many researchers worldwide. Ebrahim et al. [10] conducted experiments using the Wisconsin Diagnosis Breast Cancer database to classify breast cancer as either benign or malignant. The study employed supervised learning algorithms such as Support Vector Machines with kernels like Linear and Neural Network (NN) for comparison to accomplish

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

this task. The authors implemented the Neural Network and Support Vector Machine (SVM) approach to classify breast cancer as benign or malignant. The models' performances were analysed, where the Neural Network approach provided higher accuracy and precision compared to SVM in breast cancer classification and appeared to be a fast and efficient method. The study found that the Neural Network technique is more effective than the SVM technique in breast cancer detection.

Bhise, et al. [11] conducted a study on using Machine Learning models for performing automated diagnosis of breast cancer. The study employed CNN as a classifier model and Recursive Feature Elimination (RFE) for feature selection. Additionally, the paper compared five algorithms, namely SVM, Random Forest, KNN, Logistic Regression, and Naïve Bayes classifier. The experimentation was carried out using the BreaKHis 400X Dataset, and the system's performance was measured based on accuracy and precision. The study used activation functions such as ReLu to predict the outcomes in terms of probabilities. The results showed that CNN outperformed existing methods in terms of accuracy, precision, and size of the dataset.

Hussain, et al. [12] conducted a study on the application of machine learning techniques for automated breast cancer detection. The authors used various feature extraction strategies and compared their performance to achieve precise and efficient detection of breast cancer. The results of the study indicated that the proposed method achieved an accuracy of 96.8%, sensitivity of 97.4%, and specificity of 96.4% in detecting breast cancer. The research demonstrated the effectiveness of machine learning techniques in improving the accuracy and efficiency of breast cancer detection, which can ultimately lead to better diagnosis and treatment of the disease.

3. Dataset

Dataset for this research purpose has been collected from the Kaggle Repository which is publicly available. The detailed summary of the dataset is shown below:

Table 1: Dataset Information

SI. No.	Dataset Name	Sources	Attribute Type	Number of Attributes	Number of Instances
1	Wisconsin Diagnostic Breast Cancer(WDBC)	Kaggle Repository	Numeric and binary	32	569

- There are 569 samples of malignant and benign tumour cells in the dataset.
- The first two columns in the dataset hold the samples unique Id numbers and the matching diagnosis (M=malignant, B=benign).
- The columns 3-32 contain 30 real-value characteristics computed from digitised pictures of cell nuclei that may be used to develop a model that predicts

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

whether a tumour is benign or malignant.

- 1= Malignant(Cancerous)
- 0= Benign(Not Cancerous)

List of Attributes:

For each cell nucleus, ten real-valued characteristics are computed:
• the radius (mean of distances from centre to points on the perimeter)
• the texture (standard deviation of grey-scale values)
• perimeter
• area
• smoothness (local variation in radius lengths)
• compactness (perimeter^2 / area - 1.0)
• curvature (severity of concave portions of the contour)
• concave edges (number of concave portions of the contour)
• symmetry
• the fractal dimension ("coastline approximation" - 1)

For each picture, the mean, standard error, and "worst" or largest (mean of the three largest values) of these characteristics were calculated, yielding 30 features. For example, in dataset field 3 represents Mean Radius, field 13 represents Radius SE, and field 23 represents Worst Radius.

Class Distribution: 357 benign ,212 malignant

4. Proposed Methodology

Figure 2 depicts the phases in the proposed workflow, which include data pre-processing, training, and testing with specified models, results evaluation, and prediction of Breast Cancer. Python3 is used to carry out this task.

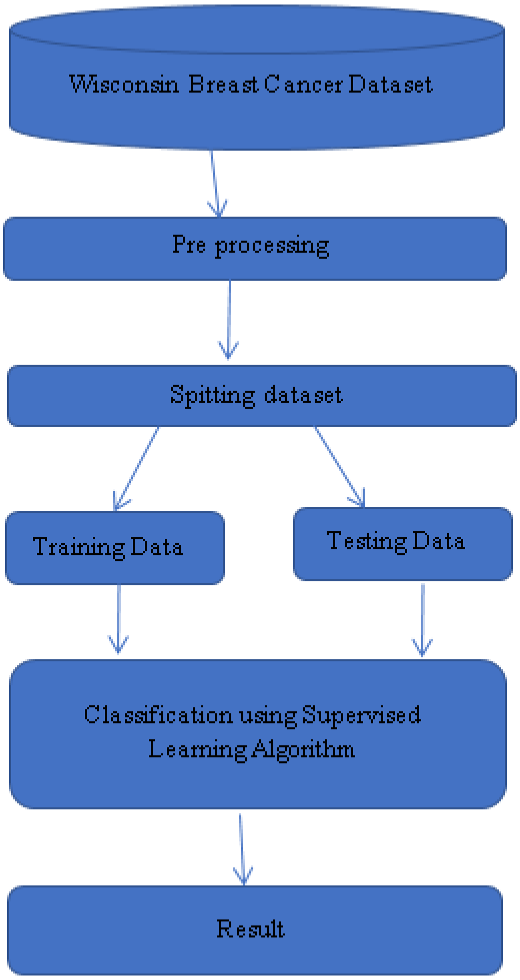


Fig-2: System Architecture

4.1. Data Pre-Processing

In Machine Learning, data pre-processing is a critical data mining technique that turns raw data into an interpretable and legible format. The first stage in constructing a Machine Learning model is data pre-processing. Real-world data is frequently partial, inconsistent, and erroneous, with mistakes or outliers and a lack of exact attribute values or patterns. Obtaining the dataset, importing libraries, importing datasets, finding missing data, encoding categorical data, partitioning the dataset into training and test sets, and feature scaling are the seven phases involved in data pre-processing. Data pre-processing is

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

essential for assuring data quality and improving the performance of Machine Learning models.

4.2. Training and Testing Model

The entire dataset has been divided into two sections, one for training and the other for testing, with a ratio of 80:20. Figure 3 depicts the final training, testing, and validation sets used for classification.

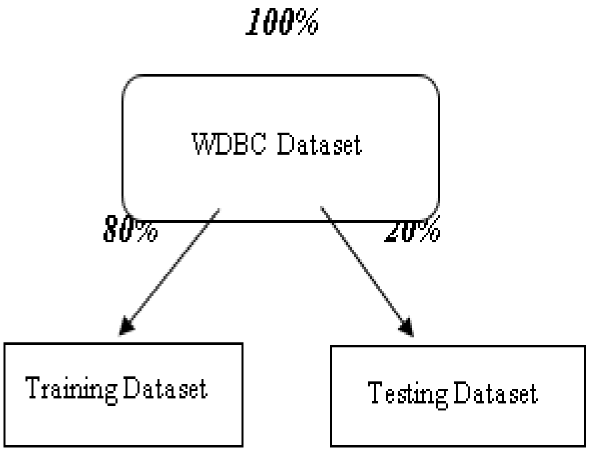


Fig-3: Final Training, Testing and Validation Sets

4.2.1. Gradient Boosting

Gradient boosting is a widely used machine learning paradigm for problems such as regression and classification. The technique generates a prediction model using an ensemble of weak prediction models, that frequently include decision trees. When a decision tree is employed as a weak learner, the resultant method is known as gradient-boosted trees, and it frequently beats random forest. Gradient-boosted tree models, like other boosting methods, are built in stages. It differs from prior boosting methods in that it allows for the optimisation of an arbitrary differentiable loss function. Gradient boosting is very adaptable and effective in a wide range of machine learning tasks because of this feature.

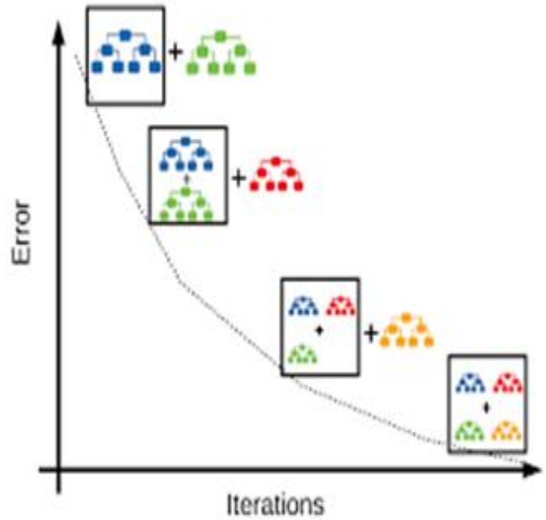


Fig-4: Gradient Boosting Classifier

4.2.2. Extreme Gradient Boosting(XGBoost)

XGBoost is a powerful Gradient Boosted decision tree framework that generates decision trees sequentially. Weights are crucial in XGBoost since they are assigned to all predictor variables and input into the decision tree to determine occurrences. The technique then increases the weight of the variables that the tree predicted erroneously and feeds them to the second decision tree. Individual classifiers/predictors combine to create a robust and precise model that can solve a wide range of problems such as regression, classification, ranking, and user-defined prediction. Trees are contributed to the ensemble one at a time and fit to correct the prediction errors generated by past models.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

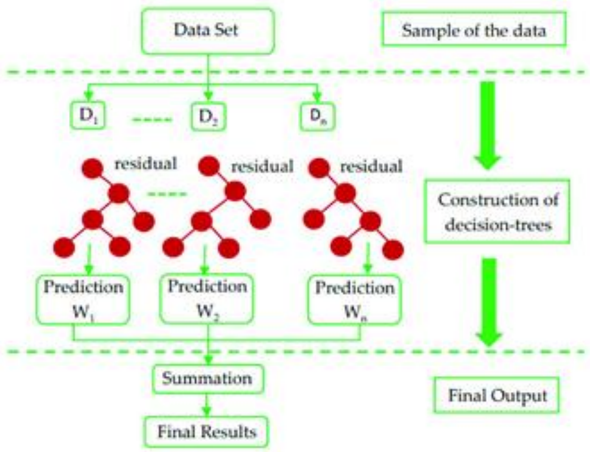


Fig-5: Extreme Boosting Classifier

4.2.3. Light Gradient Boosting(Light GBM)

Light Gradient Boosting is a gradient boosting framework that employs tree-based learning methods to enhance model efficiency while utilizing minimal memory. The framework is decision tree-based and employs two innovative techniques: gradient-based one-side sampling and exclusive feature bundling (EFB). These methods address the limitations of histogram-based algorithms, which are commonly employed in all Gradient Boosting Decision Tree (GBDT) frameworks. Gradient-based One Side Sampling and EFB enhance model training speed and accuracy while utilizing less memory. Light Gradient Boosting(Light GBM) therefore serves as a tremendously effective and productive framework for a wide range of machine learning applications.

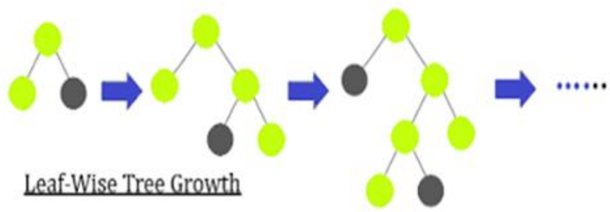


Fig-6: Light GBM Classifier

4.2.4. Decision Tree Classifier

Decision Trees constitute a form of Supervised Machine Learning in which information is constantly segmented according to specific parameters. The structure of the tree is composed of two components: decision nodes and leaves. Decision nodes indicate data split points, while leaves represent final outcomes or decisions. The tree is constructed by recursively splitting data into smaller subsets according to pertinent attributes until the appropriate degree of granularity is achieved. Due to their interpretability and ability to handle both category and numerical data, Decision Trees are widely deployed in several applications, including classification and regression tasks.

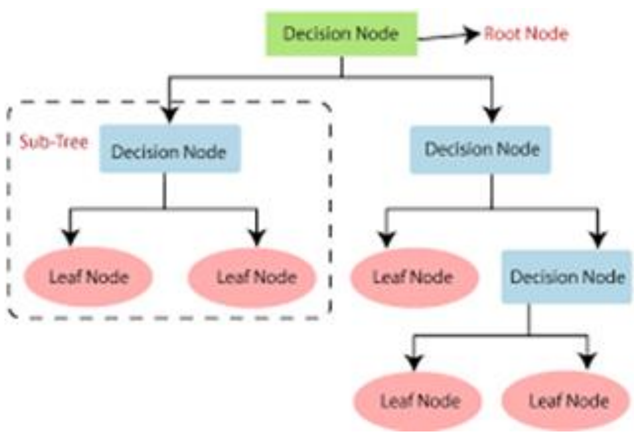


Fig-7: Decision Tree Classifier

4.2.5. K-Nearest Neighbors

K-Nearest Neighbour is a supervised learning strategy that is also one of the most basic Machine Learning algorithms. The K-NN model implies that the new data is equivalent to the known information and allocates it to the category that is nearest to the existing categories. The algorithm saves all existing information and classifies new data points according to their similarity. The K-NN methodology may be used for both Regression and Classification challenges, although it is most commonly applied for Classification. It is a non-parametric method, which indicates that no assumptions are made about the underlying data. Because it fails to learn from the training set immediately, K-NN is also characterized as a lazy learner algorithm. Instead, it captures the dataset and performs the operation.

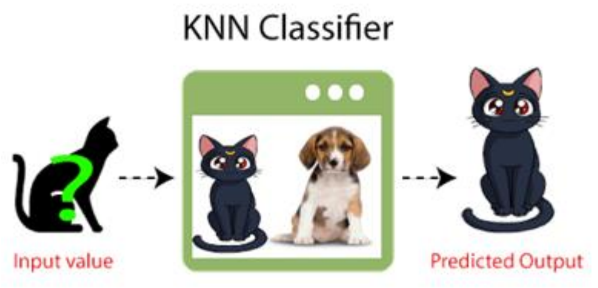


Fig-8: K-Nearest Neighbors

5. Experiments and Result:

This section goes through the Breast Cancer dataset, experiments, and evaluation scheme. Several methods for identification and analysis of data were implemented in this.

A. Experimental Setup

This section provides the parameters and examines the outcomes of the machine learning algorithms used.

Accuracy:

The accuracy of detection is often stated as a percentage of the ratio of successfully detected cases to the total number of occurrences in the dataset. It is crucial to remember, however, that the accuracy of a classifier is heavily impacted by the threshold used, which may differ between tests. As a result, comparing various classifiers based just on accuracy may not be the most practical strategy. Yet, it provides an overall picture of the classifier's effectiveness in recognising cases.

The formula for calculating accuracy is as follows:

$$Accuracy = TP + TN /(TP + TN + FP + FN)$$

Where: TP = True positive; FN= False negative; FP= False positive; TN = True negative. Similarly, P and N represent the Positive and Negative population of Malignant and Benign cases.

Recall:

The fraction of real positive observations that are accurately labelled as positive is referred to as recall or sensitivity. This statistic is especially important in the medical domain since it demonstrates the precision with which observations are diagnosed.

Mathematically, sensitivity or true positive rate (TPR) is calculated as follows: TP / (TP+FN)

While the specificity or the true negative rate (TNR) is defined by : TN / (TN + FP)

Precision:

The precision metric is used to solve the Accuracy limitation. The precision determines the fraction of valid positive predictions. It may be measured as the proportion of true positive forecasts to total positive predictions (True Positive and False Positive).

Precision = TP / (TP + TN).

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

F-Scores:

The F-score, also known as the F1 Score, is a metric used to evaluate a binary classification model based on positive class predictions. Precision and Recall are used to compute it. It is a form of composite score that includes both Precision and Recall. As a result, the F1 Score may be computed as the harmonic mean of accuracy and recall, with equal weights assigned to each.

The formula for calculating the F1 score is given below: 2*(precision*recall)/(precision+recall)

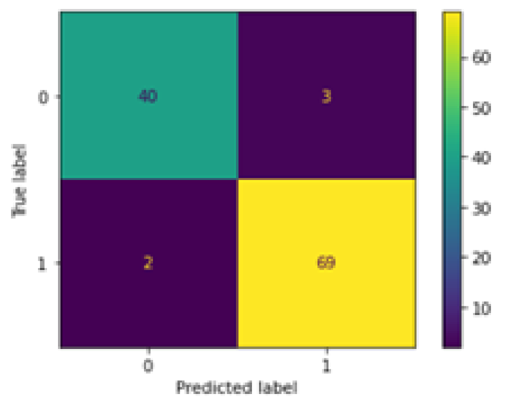
Table 2: Overall Results for Wisconsin Diagnostic Breast Cancer Data Using

Algorithms	Precision Score	Recall Score	F1-Score	Accuracy
Gradient boosting	92.83	95.18	93.50	92.61
Light Gradient Boosting	99.12	99.59	99.22	98.97
Extreme Gradient Boosting	92.83	95.18	93.50	92.61
Decision Tree	91.44	92.77	92.10	90.86
K-Nearest Neighbors	89.65	90.65	88.29	86.98

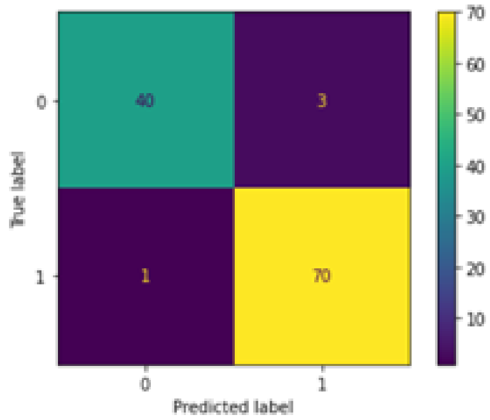
Machine Learning Algorithms

Evaluation of various machine learning models on WDBC dataset observed an accuracy in the range of (86.98% to 98.97 %) on the original dataset. KNN has produced the least accuracy of 86.98%. The confusion matrixes of all Machine Learning algorithms also describe the results of the prediction model.

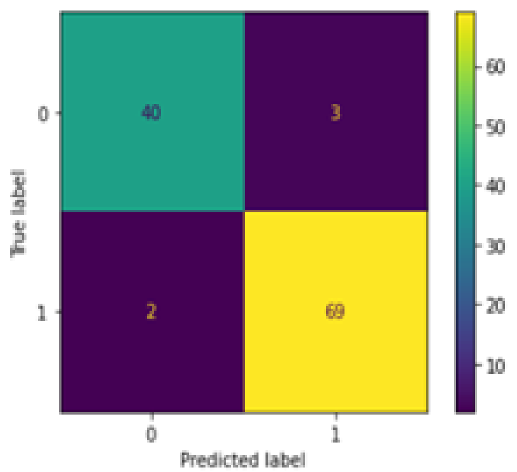
Confusion Matrix of Gradient Boosting



Confusion Matrix of Light Gradient Boosting

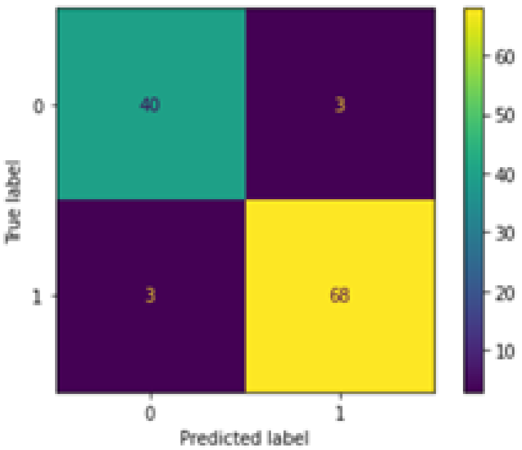


Confusion Matrix of Extreme Gradient Boosting



ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Confusion Matrix of Decision Tree



Confusion Matrix of K-Nearest Neighbors

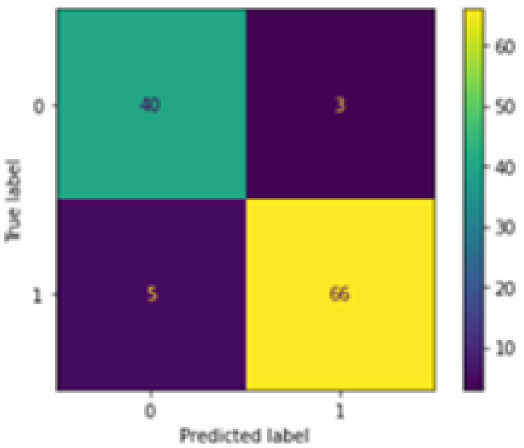
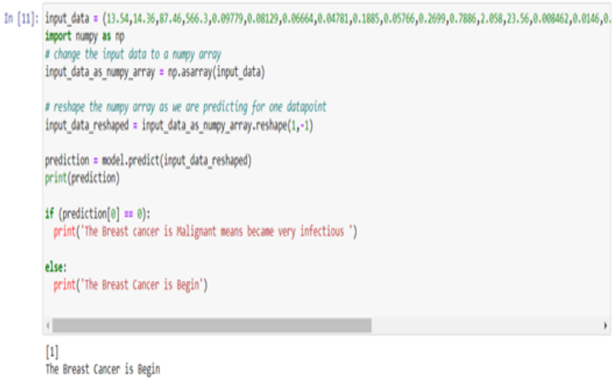


Fig-9: Confusion Matrices for different Machine Learning Algorithms



```
In [11]: input_data = (13.54,14.36,87.46,566.3,0.09779,0.08129,0.06664,0.04781,0.1885,0.05766,0.2699,0.7886,2.058,23.56,0.008462,0.0146,0.
import numpy as np
# change the input data to a numpy array
input_data_as_numpy_array = np.asarray(input_data)

# reshape the numpy array as we are predicting for one datapoint
input_data_reshaped = input_data_as_numpy_array.reshape(1,-1)

prediction = model.predict(input_data_reshaped)
print(prediction)

if (prediction[0] == 0):
  print('The Breast cancer is Malignant means became very infectious ')

else:
  print('The Breast Cancer is Begin')

[1]
The Breast Cancer is Begin
```

Fig-10: Classification of Breast Cancer

6. Conclusion

This section describes the project's last stages, which included testing multiple machine learning approaches for detecting benign and malignant tumours in breast cancer samples. To examine the performance of the models constructed for Breast cancer detection on the WDBC Dataset, various performance evaluation measures were utilised. KNN, Decision Tree, Gradient Boost, Extreme Gradient Boost, and Light Gradient Boost were among the algorithms evaluated. After a comparison, it was determined that Light GBM was the most accurate and exact approach. Light GBM functioned admirably, achieving 98.97% accuracy. Utilizing the information acquired, the system can determine if a patient has benign or malignant breast cancer. Overall, the proposed strategy is successful and efficient at detecting breast cancer.

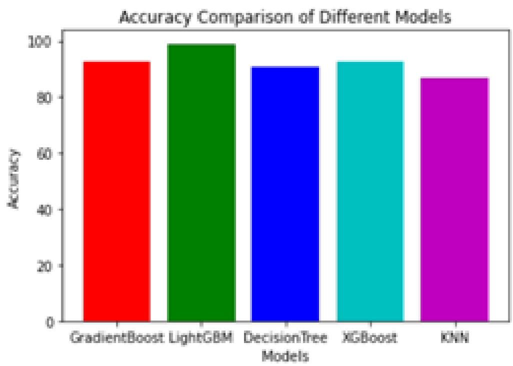


Fig-11: Testing Accuracy comparision of Various ML Algorithms

7. References

[1] Sharma, S., Aggarwal, A., & Choudhury, T. (2018, December). Breast cancer detection using machine learning algorithms. In 2018 International conference on computational techniques, electronics and mechanical systems (CTEMS) (pp. 114-118). IEEE.

[2] Mohammed, S. A., Darrab, S., Noaman, S. A., & Saake, G. (2020). Analysis of breast cancer detection using different machine learning techniques. In Data Mining and Big Data: 5th International Conference, DMBD 2020, Belgrade, Serbia, July 14–20, 2020, Proceedings 5 (pp. 108-117). Springer Singapore.

[3] Tahmooresi, M., Afshar, A., Rad, B. B., Nowshath, K. B., & Bamiah, M. A. (2018). Early detection of breast cancer using machine learning techniques. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(3-2), 21-27.

[4] Khourdifi, Y., & Bahaj, M. (2018, December). Applying best machine learning algorithms for breast cancer prediction and classification. In 2018 International conference on electronics, control, optimization and computer science (ICECOCS) (pp. 1-5). IEEE.

[5] Bayrak, E. A., Kırcı, P., & Ensari, T. (2019, April). Comparison of machine learning methods for breast cancer diagnosis. In 2019 Scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT) (pp. 1-3). IEEE.

[6] Nallamala, S. H., Pathuri, S. K., & Koneru, S. V. (2018). A literature survey on data mining approach to effectively handle cancer treatment. International Journal of Engineering and Technology (UAE), 7, 729-732.

[7] Alarabeyyat, A., & Alhanahnah, M. (2016, August). Breast cancer detection using k-nearest neighbor machine learning algorithm. In 2016 9th International Conference on Developments in eSystems Engineering (DeSE) (pp. 35-39). IEEE.

[8] 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.

[9] Sharma, A., Kulshrestha, S., & Daniel, S. (2017, December). Machine learning approaches for breast cancer diagnosis and prognosis. In 2017 International conference on soft computing and its engineering applications (icSoftComp) (pp. 1-5). IEEE.

[10] Ali, E. E. E., & Feng, W. Z. (2016). Breast cancer classification using support vector machine and neural network. International Journal of Science and Research, 5(3), 1-6.

[11] Bhise, S., Gadekar, S., Gaur, A. S., Bepari, S., & Deepmala Kale, D. S. A. (2021). Breast cancer detection using machine learning techniques. Int. J. Eng. Res. Technol, 10(7).

[12] Hussain, L., Aziz, W., Saeed, S., Rathore, S., & Rafique, M. (2018, August). Automated breast cancer detection using machine learning techniques by extracting different feature extracting strategies. In 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE) (pp. 327-331). IEEE.