2026 Volume: 02 Issue: 08 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-18

AN EXPLAINABLE MACHINE LEARNING FRAMEWORK FOR BENIGN AND MALIGNANT BREAST TUMOR CLASSIFICATION USING MULTIPLE CLASSIFICATION MODELS

***[1]Upasana Sarkar, [1]Rima Pal, [2]Tathagata Roy Chowdhury**

[1]Student, Department of Computer Science and Engineering Techno Institute of Engineering and Management Kolkata, India.

[2]Assistant Professor, Department of Computer Science and Engineering Techno Institute of Engineering and Management Kolkata, India.

Article Received: 10 July 2026
Article Revised: 30 July 2026
Published on: 19 August 2026

***Corresponding Author: Upasana Sarkar**
Student, Department of Computer Science and Engineering Techno Institute of Engineering and Management Kolkata, India.

DOI: https://doi-doi.org/101555/ijrpa.5351

ABSTRACT

Breast tumor classification is an important machine learning application in healthcare because reliable differentiation between benign and malignant observations may sup-port computer-assisted analysis. This study presents an educational machine learning framework for classifying breast tumor observations into benign and malignant categories using quantitative measurements of cell nuclei. The Breast Cancer Wisconsin (Diagnos-tic) dataset, containing 569 observations and 30 numerical predictive features, was used for experimentation. Six classification algorithms were evaluated: Logistic Regression, K-Nearest Neighbors, Decision Tree, Random Forest, Support Vector Machine, and Gra-dient Boosting. The experimental workflow consisted of feature preparation, train-test partitioning, model training, prediction, and multi-metric evaluation. Accuracy, pre-cision, recall, F1-score, ROC-AUC, and confusion matrices were used for comparative assessment. In the reported held-out test split, Random Forest achieved approximately 97.37% accuracy, 100.00% precision, 92.86% recall, 96.30% F1-score, and 99.8% ROC-AUC. Feature-importance analysis indicated that perimeter-worst, area-worst, concave-points-mean, concave-points-worst, and radius-worst were among the most influential variables in the Random Forest model. An interactive prediction interface was also de-veloped to demonstrate model inference. The system is intended for educational and research purposes and must not be interpreted as a clinical diagnostic instrument.

KEYWORDS: Breast Tumor Classification; Machine Learning; Random Forest; Breast Cancer Wiscon-sin Dataset; Explainable AI; Binary Classification.

INTRODUCTION

Breast cancer is a major global health problem. The World Health Organization reported an estimated 2.4 million women diagnosed with breast cancer and approximately 694,000 deaths globally in 2024 [1]. Reliable and timely diagnosis is therefore an important component of breast cancer care.

Machine learning has increasingly been investigated for computer-assisted analysis of biomedical data. In breast cancer research, algorithms can learn relationships between quantitative measurements and diagnostic classes, potentially providing useful decision-support or educational tools. However, predictive performance on a benchmark dataset should not be confused with clinical diagnostic effectiveness.

The present study uses the Breast Cancer Wisconsin (Diagnostic) (WDBC) dataset. The UCI Machine Learning Repository describes the dataset as a classification dataset con-taining 569 observations and 30 real-valued features derived from digitized fine-needle as-pirate (FNA) images of breast masses [2]. The original feature-extraction work described measurements related to nuclear size, shape, and texture, including mean, standard error, and worst-value summaries [3, 4].

The WDBC dataset has become a widely used benchmark for evaluating classification methods. Previous studies have investigated support vector machines, logistic regression, nearest-neighbour methods, decision trees, ensemble models, and other machine learning techniques on the dataset [5, 6, 7, 8, 9].

The research problem addressed in this study is the comparative evaluation of multiple machine learning algorithms under a common experimental workflow and the integration of the selected model into an interactive prediction interface.

The objectives are:

To develop a reproducible machine learning workflow for benign and malignant breast tumor classification.

To compare six supervised classification algorithms.

To evaluate models using accuracy, precision, recall, F1-score, ROC-AUC, and con-fusion matrices.

To identify influential features using Random Forest feature importance.

To select a strong-performing model for an interactive educational prediction ap-plication.

To demonstrate the complete workflow from dataset preparation to model inference and interpretation.

The study follows established principles of supervised statistical learning and model evaluation [16, 17].

MATERIALS AND METHODS

Dataset

The Breast Cancer Wisconsin (Diagnostic) dataset was obtained from the UCI Machine Learning Repository. It contains 569 observations, 30 numerical features, and a binary diagnosis target [2]. The feature values were computed from digitized images of fine-needle aspirates of breast masses and describe characteristics of cell nuclei [3, 4].

The dataset contains 212 malignant observations and 357 benign observations [4]. The class distribution is therefore not exactly balanced, and this characteristic was retained rather than artificially changing the original class proportions.

Table 1: Summary of the WDBC dataset used in the study.

Property	Value
Total observations	569
Predictive numerical features	30
Diagnostic classes	2
Malignant observations	212
Benign observations	357
Learning task	Binary classification
Feature type	Continuous

Feature Groups

The 30 predictive variables are organized into ten basic nuclear characteristics:

1. Radius,
2. Texture,
3. Perimeter,
4. Area,
5. Smoothness,
6. Compactness,
7. Concavity,

8. Concave points,
9. Symmetry, and
10. Fractal dimension.

Each characteristic is represented through mean, standard-error, and worst-value measurements. This structure follows the original feature extraction process reported for the dataset [3].

Software and Implementation

The experimental workflow was implemented in Python using standard scientific com-puting and machine learning tools. Scikit-learn provides implementations of the classi-fication algorithms and evaluation procedures used in the workflow [18]. Data handling and numerical processing can be performed using commonly adopted Python scientific-computing libraries [30, 31].

The workflow can be represented as:

Dataset → Feature/Target Separation → Preprocessing → Train/Test Split → Model Training → Prediction → Evaluation → Model Selection Interactive Application

Data Preprocessing

The diagnosis column was treated as the target variable and the numerical measurements as predictors. For models that are sensitive to feature scale, standardized numerical features were used. Scaling and preprocessing are important because algorithms such as KNN, Logistic Regression, and SVM can be influenced by differences in feature magnitude [16, 17].

The reported confusion matrices contain 114 test observations, which is consistent with an 80:20 held-out partition of the 569-observation dataset after rounding.

Classification Models

Six supervised classification algorithms were evaluated.

Logistic Regression: Logistic Regression was used as a linear probabilistic baseline for binary classification.

K-Nearest Neighbors: KNN predicts a class according to nearby training observa-tions. Its performance can depend on the distance measure and the selected number of neighbours [14, 32].

Decision Tree: Decision-tree learning recursively partitions observations according to feature-based rules. The CART family of tree methods provides a standard foundation for tree-based classification [12].

Random Forest: Random Forest combines multiple randomized decision trees and aggregates their predictions. The method was introduced by Breiman and is widely used for classification and regression [11].

Support Vector Machine: SVM constructs a separating decision boundary based on margin maximization and kernel methods [13].

Gradient Boosting: Gradient Boosting constructs an additive sequence of weak learners to improve predictive performance by optimizing a differentiable loss function [15].

These models represent different learning principles, allowing the study to compare linear, neighbourhood-based, tree-based, ensemble, kernel-based, and boosting approaches [16].

Evaluation Metrics

The following metrics were used.

Accuracy:

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

Precision:

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

Recall:

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

F1-score:

$$F1 = 2\frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

ROC curves and their area under the curve (AUC) were also used. ROC analysis is a standard approach for evaluating binary classifiers over varying decision thresholds [19]. Precision-recall and ROC-based evaluation can provide complementary views of classifier behaviour [21].

Confusion matrices were generated to examine the distribution of true-positive, true-negative, false-positive, and false-negative predictions.

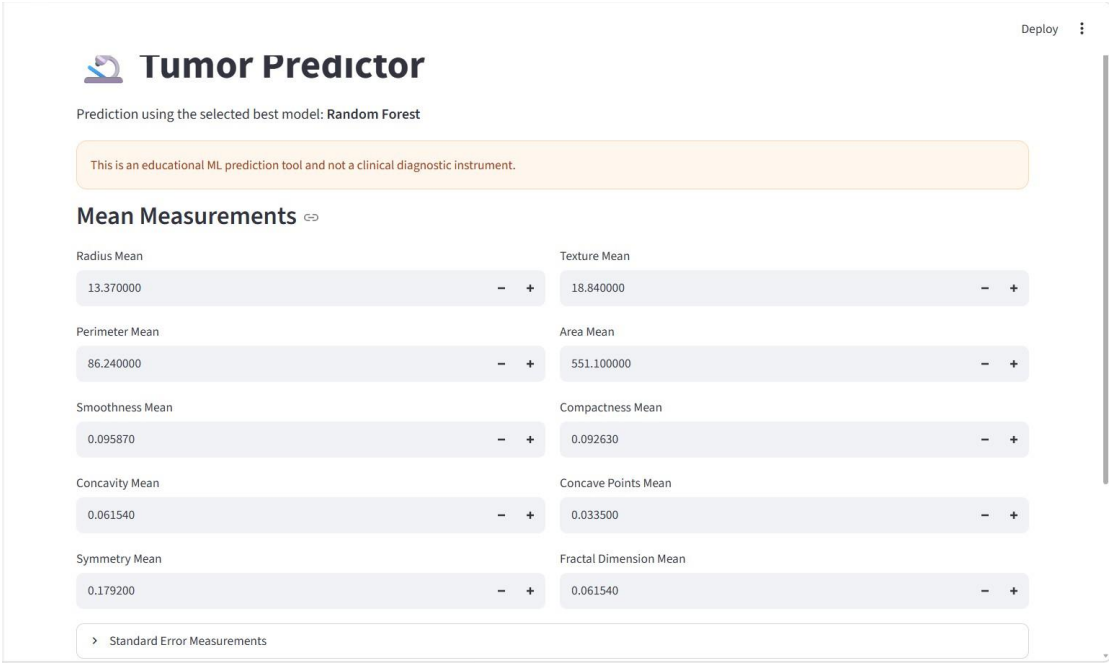

Figure 1: Interactive Tumor Predictor interface showing a benign test input using the mean-measurement fields.

Interactive Prediction Interface

An interactive prediction interface was developed to demonstrate how the selected model can be used with manually entered numerical measurements. The interface contains mean, standard-error, and worst-measurement sections. Figure 1 shows the mean-measurement portion of the interface used during testing.

The application displays the predicted class together with model probability estimates. These values represent model outputs for the supplied numerical feature vector and are not equivalent to clinically validated probabilities.

RESULTS AND DISCUSSION

Comparative Model Performance

The experimental results supplied with the project show that Random Forest and Support Vector Machine achieved the highest reported accuracy of approximately 97.37%, while Random Forest achieved the highest reported ROC-AUC of approximately 0.998.

Table 2: Reported performance of the six evaluated machine learning models.

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Random Forest	97.37%	100.00%	92.86%	96.30%	0.998
Logistic Regression	96.49%	97.50%	92.86%	95.12%	0.996
Support Vector Machine	97.37%	100.00%	92.86%	96.30%	0.995
Gradient Boosting	96.49%	100.00%	90.48%	95.00%	0.994
K-Nearest Neighbors	95.61%	97.44%	90.48%	93.83%	0.982
Decision Tree	92.11%	94.59%	83.33%	88.61%	0.945

Important reproducibility note: The values in Table 2 are transcribed from the supplied project screenshots. Before journal submission, the exact values should be re-generated directly from the final Python training/evaluation script and copied into the manuscript without rounding inconsistencies.

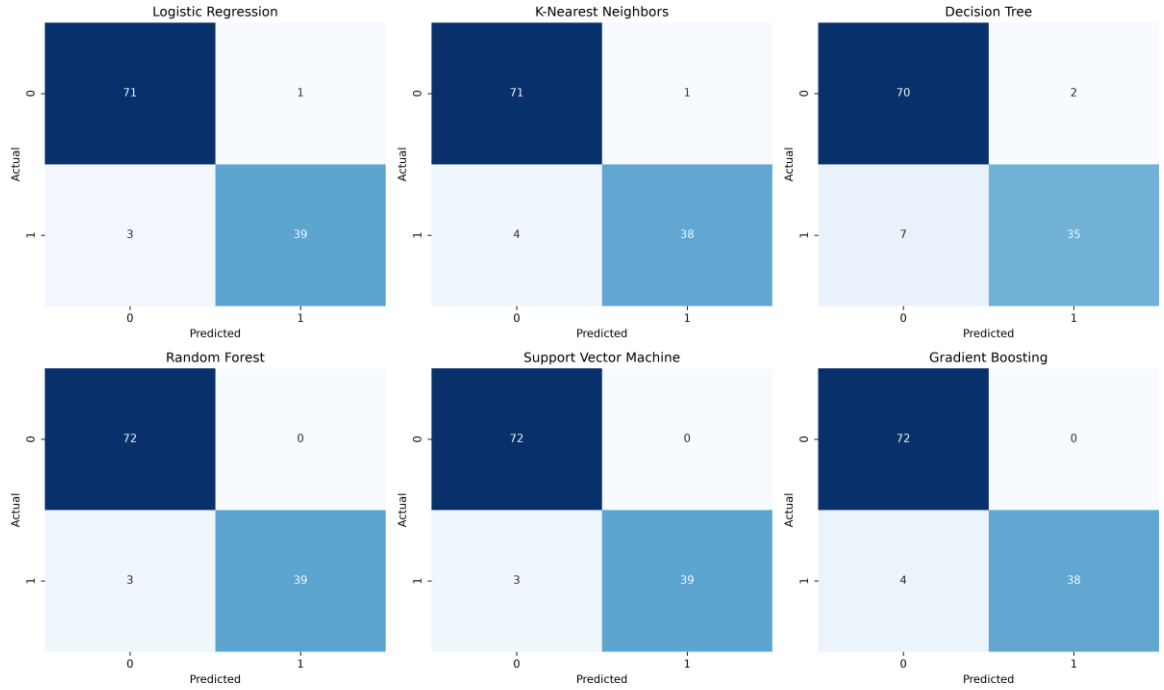

Figure 2: Confusion matrices for Logistic Regression, KNN, Decision Tree, Random Forest, SVM, and Gradient Boosting.

Confusion Matrix Analysis

The Random Forest confusion matrix contains 72 correctly classified benign observations, 39 correctly classified malignant observations, 3 malignant observations incorrectly classified as benign, and no benign observations incorrectly classified as malignant.

Thus:

$$TN = 72,$$

$$FP = 0,$$

$$FN = 3,$$

$$TP = 39.$$

The resulting accuracy is:

$$\frac{72 + 39}{72 + 0 + 3 + 39} = 0.9737, \qquad (5)$$

or approximately 97.37%.

The three false-negative predictions are particularly important from a research perspec-tive because they demonstrate that high overall accuracy does not mean perfect classifi-cation. Medical machine learning studies should therefore examine sensitivity/recall and false-negative behaviour rather than reporting accuracy alone [21, 22].

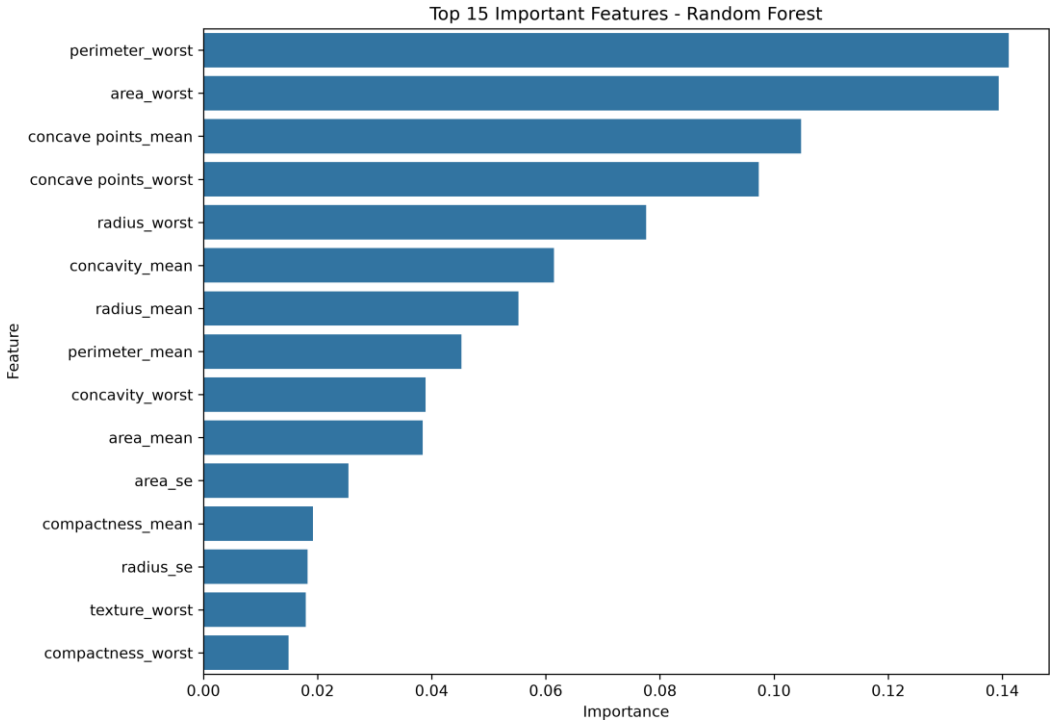

Figure 3: Top 15 feature-importance scores obtained from the Random Forest model.

Feature Importance

Figure 3 shows the top 15 features ranked by the Random Forest feature-importance measure. The leading variables in the displayed experiment were:

1. Perimeter-worst,
2. Area-worst,

3. Concave-points-mean,
4. Concave-points-worst,
5. Radius-worst,
6. Concavity-mean,
7. Radius-mean,
8. Perimeter-mean,
9. Concavity-worst,
10. Area-mean,
11. Area-se,
12. Compactness-mean,
13. Radius-se,
14. Texture-worst,
15. Compactness-worst.

The displayed ranking suggests that measurements describing the size and shape of cell nuclei contributed strongly to the Random Forest decision process. This is consistent with the original WDBC feature construction, which was explicitly designed around nuclear size, shape, and texture characteristics [3, 4].

Feature importance should not, however, be interpreted as a causal medical explanation. A feature-importance score indicates how a trained model uses variables within the fitted predictive system; it does not establish that an individual variable independently causes malignancy.

For future work, model-agnostic explanation techniques such as LIME and SHAP could be incorporated to provide local explanations for individual predictions [25, 26]. Recent reviews indicate that SHAP has become a frequently used XAI approach in breast cancer research, particularly for tree-based models [27].

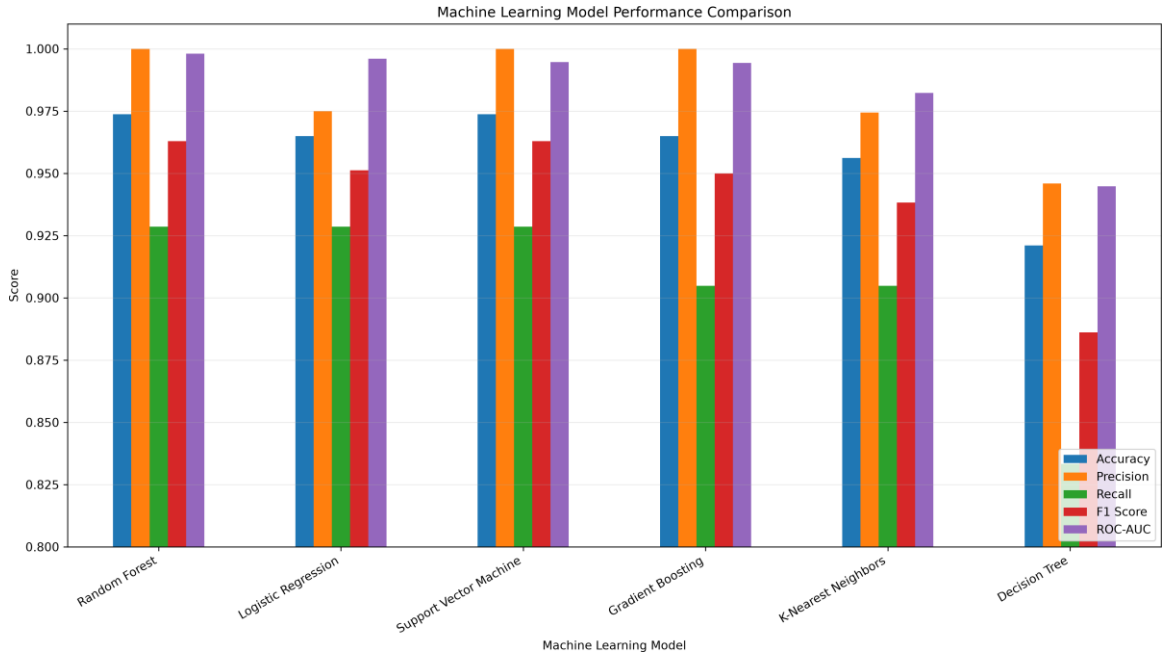

Figure 4: Comparative bar chart of accuracy, precision, recall, F1-score, and ROC-AUC for the six models.

Performance Comparison

Figure 4 illustrates the multi-metric comparison. Random Forest provided the strongest combination of accuracy and ROC-AUC in the reported experiment. Logistic Regression and SVM also performed strongly, demonstrating that the dataset is highly separable using several conventional algorithms.

Previous WDBC studies similarly report strong performance for SVM, Logistic Regres-sion, KNN, Decision Tree, and Random Forest, although reported values vary according to the train-test partition, preprocessing, feature selection, and hyperparameter configu-ration [5, 6, 7, 8, 9]. Therefore, the numerical results of different studies should not be compared as though they were generated under identical experimental conditions.

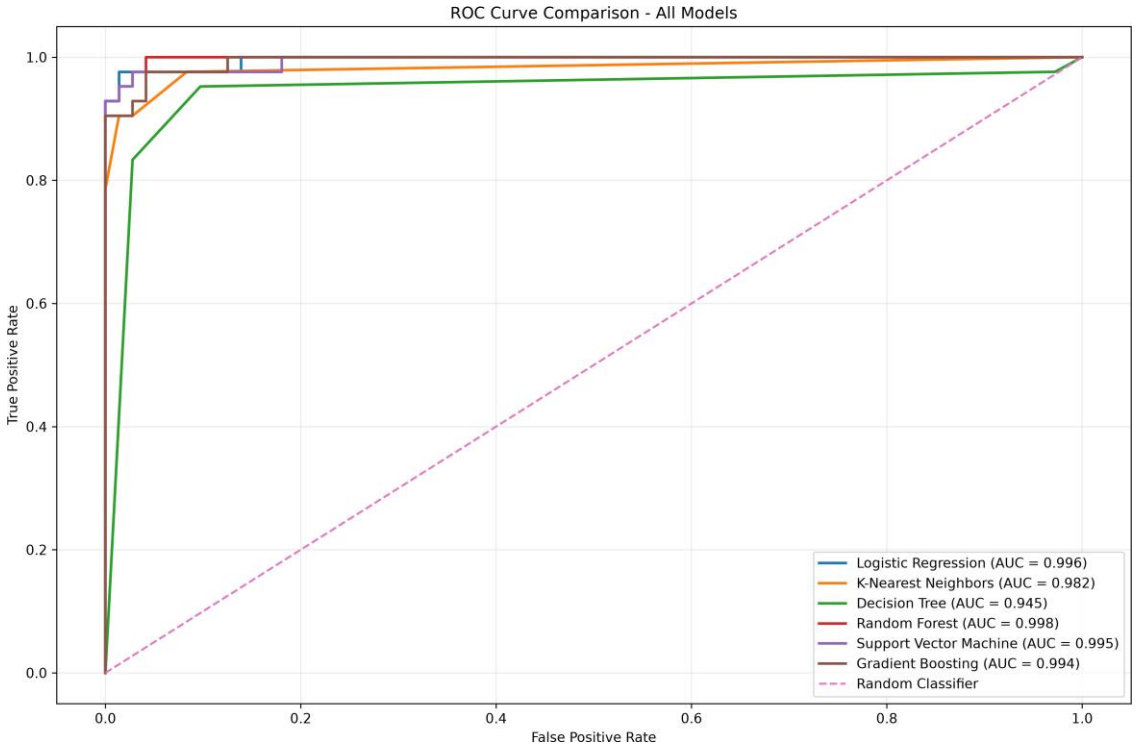

Figure 5: ROC curve comparison for the six evaluated classification models.

ROC-AUC Analysis

Figure 5 shows the ROC curves generated by the project. The displayed AUC values were approximately 0.996 for Logistic Regression, 0.982 for KNN, 0.945 for Decision Tree, 0.998 for Random Forest, 0.995 for SVM, and 0.994 for Gradient Boosting.

The Random Forest curve produced the largest reported AUC, indicating strong discrimination between the two classes in the held-out test experiment. ROC-AUC is useful because it evaluates ranking performance across multiple thresholds rather than relying on a single classification threshold [19].

Nevertheless, a high AUC on a benchmark dataset does not establish clinical validity. External validation, calibration, prospective testing, and assessment across representative populations would be required for clinical use [24, 23].

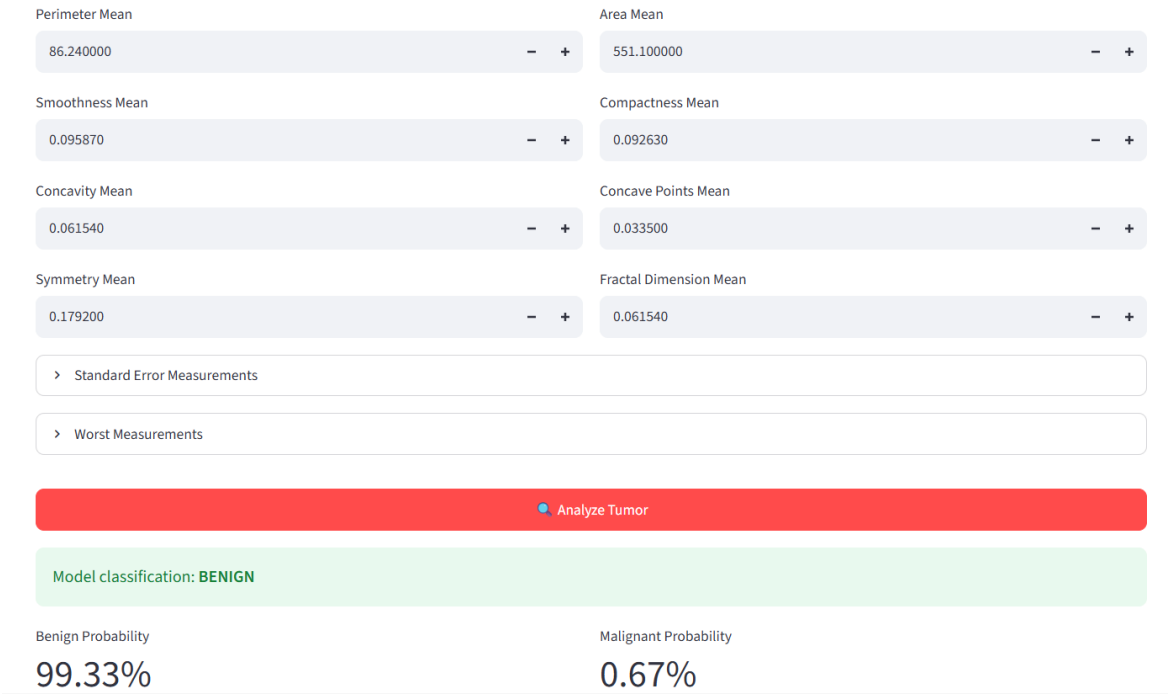

Figure 6: Interactive application output for the benign test example. The model classified the supplied feature vector as BENIGN with a displayed benign probability of 99.33%.

Benign Test Example

The supplied application screenshot shows a test configuration that was classified as BENIGN. The interface displayed:

Model classification: BENIGN

Benign probability: 99.33%

Malignant probability: 0.67%

This example demonstrates the prediction interface rather than a clinical patient case. The displayed probability is a model output and should not be interpreted as a clinically calibrated probability of disease.

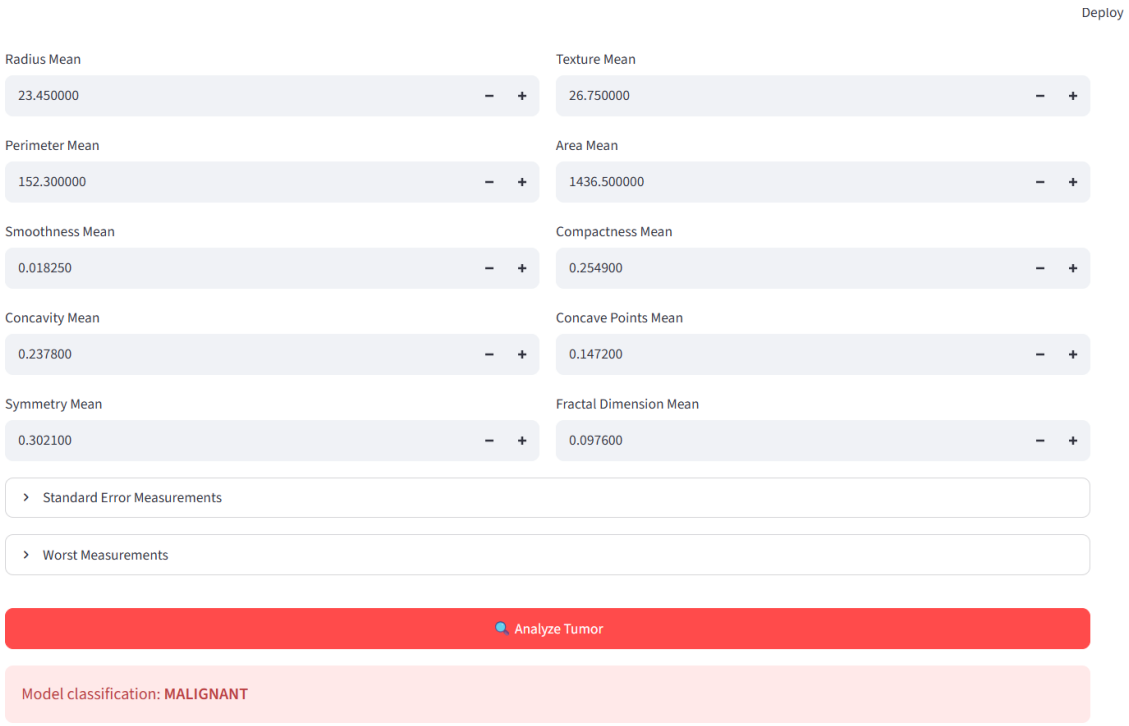

Figure 7: Interactive application showing the malignant test configuration and the resulting MALIGNANT classification.

Malignant Test Example

A second manually entered feature configuration was evaluated to verify that the appli-cation could also produce a malignant classification. The displayed mean measurements were:

Table 3: Mean measurements shown in the malignant test example.

Feature	Value
Radius Mean	23.450000
Texture Mean	26.750000
Perimeter Mean	152.300000
Area Mean	1436.500000
Smoothness Mean	0.018250
Compactness Mean	0.254900
Concavity Mean	0.237800
Concave Points Mean	0.147200
Symmetry Mean	0.302100
Fractal Dimension Mean	0.097600

The application returned:

Model classification: MALIGNANT

This test demonstrates application functionality. It should not be presented as an inde-pendent clinical validation sample.

Interpretation and Comparison with Existing Work

The experimental results are broadly consistent with previous studies showing that conventional supervised learning algorithms can achieve high classification performance on the WDBC dataset. Agarap evaluated several machine learning algorithms and reported test accuracies above 90% for the presented models [5]. Zheng, Yoon, and Lam re-ported approximately 97.38% accuracy for a hybrid K-means/SVM method on WDBC [6]. Wang et al. investigated an ensemble of SVM models for breast cancer diagnosis and demonstrated the potential value of combining multiple classifiers [7]. Rajaguru and Chakravarthy compared Decision Tree and KNN methods on WDBC after feature selection [8].

Recent work continues to compare machine learning algorithms on WDBC and related breast cancer datasets, confirming that performance depends strongly on preprocessing, feature selection, model configuration, and evaluation protocol [9, 10]. This supports the use of a multi-model comparison rather than selecting a classifier solely because it is popular.

The present project also emphasizes interpretability. Tree-based feature importance pro-vides a global view of the variables used by Random Forest, while methods such as LIME and SHAP can provide local explanations for individual predictions [25, 26]. Explain-ability is especially relevant in healthcare because predictive performance alone does not guarantee that a model is understandable, trustworthy, or clinically appropriate [27, 28].

CONCLUSION

This study developed a machine learning framework for benign and malignant breast tumor classification using the Breast Cancer Wisconsin (Diagnostic) dataset. Six classifi-cation algorithms were compared using accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrices.

In the reported held-out experiment, Random Forest achieved approximately 97.37% accuracy and the highest reported ROC-AUC of approximately 0.998. The Random Forest confusion matrix contained 72 true-negative, 39 true-positive, 0 false-positive, and 3 false-negative predictions. Feature-importance analysis identified perimeter-worst, area-worst, concave-points-mean, concave-points-worst, and radius-worst among the leading variables.

The project also demonstrates the integration of the selected model into an interactive prediction interface. This provides a complete educational machine learning workflow from dataset preparation through model comparison, feature analysis, and inference.

However, the study has important limitations. The WDBC dataset is a relatively small benchmark dataset, the evaluation shown here is based on a single held-out test split, and the

interactive probability outputs have not been independently calibrated for clinical use. The reported results should therefore not be interpreted as evidence that the system can diagnose breast cancer.

Future research should include repeated stratified cross-validation, independent external validation, probability calibration, systematic hyperparameter tuning, uncertainty anal-ysis, fairness evaluation, and more detailed XAI methods such as SHAP and LIME. Clin-ical deployment would additionally require appropriate clinical validation, governance, regulatory assessment, and prospective evaluation.

ACKNOWLEDGEMENTS

The author acknowledges the academic environment and faculty support of Techno Engi-neering College Banipur, affiliated with Maulana Abul Kalam Azad University of Technol-ogy (MAKAUT), West Bengal, during the development of this student research project. The author also acknowledges the UCI Machine Learning Repository for providing access to the Breast Cancer Wisconsin (Diagnostic) dataset [2].

REFERENCES

1. World Health Organization, “Breast cancer,” WHO Fact Sheet, 3 July 2026. Avail-able: https://www.who.int/news-room/fact-sheets/detail/breast-cancer
2. W. H. Wolberg, W. N. Street, and O. L. Mangasarian, “Breast Cancer Wis-consin (Diagnostic),” UCI Machine Learning Repository, Dataset 17, 1995. DOI: 10.24432/C5DW2B.
3. W. N. Street, W. H. Wolberg, and O. L. Mangasarian, “Nuclear feature extrac-tion for breast tumor diagnosis,” in Biomedical Image Processing and Biomedical Visualization, Proc. SPIE, vol. 1905, pp. 861–870, 1993. DOI: 10.1117/12.148698.
4. W. H. Wolberg, W. N. Street, D. M. Heisey, and O. L. Mangasarian, “Computer-derived nuclear features distinguish malignant from benign breast cytology,” Human Pathology, vol. 26, no. 7, pp. 792–796, 1995. DOI: 10.1016/0046-8177(95)90229-5.
5. F. Agarap, “On breast cancer detection: An application of machine learning algorithms on the Wisconsin Diagnostic Dataset,” arXiv:1711.07831, 2018. DOI: 10.48550/arXiv.1711.07831.
6. Zheng, S. W. Yoon, and S. S. Lam, “Breast cancer diagnosis based on fea-ture extraction using a hybrid of K-means and support vector machine algo-rithms,” Expert Systems with Applications, vol. 41, no. 4, pp. 1476–1482, 2014. DOI: 10.1016/j.eswa.2013.08.044.
7. H. Wang, B. Zheng, S. W. Yoon, and H. S. Ko, “A support vector machine-based ensemble

algorithm for breast cancer diagnosis," European Journal of Operational Research, vol. 267, no. 2, pp. 687–699, 2018. DOI: 10.1016/j.ejor.2017.12.001.

8. H. Rajaguru and S. C. S. R. Chakravarthy, "Analysis of Decision Tree and K-Nearest Neighbor Algorithm in the Classification of Breast Cancer," Asian Pa-cific Journal of Cancer Prevention, vol. 20, no. 12, pp. 3777–3781, 2019. DOI: 10.31557/APJCP.2019.20.12.3777.
9. M. M. Hossin, F. M. J. M. Shamrat, M. R. Bhuiyan, R. A. Hira, T. Khan, and S. Molla, "Breast cancer detection: an effective comparison of different machine learning algorithms on the Wisconsin dataset," Bulletin of Electrical Engineering and Informatics, vol. 12, no. 4, 2023.
10. Acet and A. E. Akkaya, "Performance analysis of machine learning algorithms on Wisconsin Diagnostic Breast Cancer Data Set enriched with data augmentation technique," International Journal of Applied Methods in Electronics and Computers, 2026. DOI: 10.58190/ijamec.2026.171.
11. L. Breiman, "Random forests," Machine Learning, vol. 45, pp. 5–32, 2001. DOI: 10.1023/A:1010933404324.
12. L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, Classification and Regression Trees. Belmont, CA, USA: Wadsworth, 1984.
13. C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, pp. 273–297, 1995. DOI: 10.1007/BF00994018.
14. T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21–27, 1967. DOI: 10.1109/TIT.1967.1053964.
15. J. H. Friedman, "Greedy function approximation: A gradient boosting ma-chine," The Annals of Statistics, vol. 29, no. 5, pp. 1189–1232, 2001. DOI: 10.1214/aos/1013203451.
16. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed. New York, NY, USA: Springer, 2009. DOI: 10.1007/978-0-387-84858-7.
17. G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning: With Applications in R, 2nd ed. New York, NY, USA: Springer, 2021. DOI: 10.1007/978-1-0716-1418-1.
18. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cour-napeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
19. T. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861–874, 2006. DOI: 10.1016/j.patrec.2005.10.010.

20. T. Roy Chowdhury and B. Mollah, “Artificial Intelligence Meets Healthcare Infor-matics: A Comprehensive Review of Object Recognition for Clinical Decision Sup-port and Smart Medical Systems,” International Journal of Research Publication and Reviews, vol. 2, pp. 185–197, 2026.
21. T. Saito and M. Rehmsmeier, “The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets,” PLoS ONE, vol. 10, no. 3, e0118432, 2015. DOI: 10.1371/journal.pone.0118432.
22. E. W. Steyerberg, A. Vickers, N. Cook, N. Gerds, M. Gonen, N. Obuchowski, M. Pencina, and M. Kattan, “Assessing the performance of prediction models: A frame-work for some traditional and novel measures,” Epidemiology, vol. 21, no. 1, pp. 128–138, 2010. DOI: 10.1097/EDE.0b013e3181c30fb2.
23. G. S. Collins, J. B. Reitsma, D. G. Altman, and K. G. M. Moons, “Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement,” Annals of Internal Medicine, vol. 162, no. 1, pp. 55–63, 2015. DOI: 10.7326/M14-0697.
24. C. J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, and D. King, “Key challenges for delivering clinical impact with artificial intelligence,” BMC Medicine, vol. 17, article 195, 2019. DOI: 10.1186/s12916-019-1426-2.
25. M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you? Explaining the predictions of any classifier,” in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1135–1144, 2016. DOI: 10.1145/2939672.2939778.
26. S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predic-tions,” in Advances in Neural Information Processing Systems 30, 2017.
27. Ghasemi, S. Hashtarkhani, D. L. Schwartz, and A. Shaban-Nejad, “Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review,” Cancer Innovation, 2024. DOI: 10.1002/cai2.120.
28. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean, “A guide to deep learning in healthcare,” Nature Medicine, vol. 25, pp. 24–29, 2019. DOI: 10.1038/s41591-018-0316-z.
29. S. Das and T. Roy Chowdhury, “Evaluating an AI-Assisted Mental Stress Screening Tool for Occupational Health and Nursing Staff,” Asian Journal of Research in Nurs-ing and Health, vol. 9, no. 1, pp. 1754–1761, 2026. DOI: 10.9734/ajrnh/2026/v9i1393.
30. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M.

Brett, A. Haldane, J. F. del Rıo, M. Wiebe, P. Peterson, P. G´erard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi,

31. Gohlke, and T. E. Oliphant, “Array programming with NumPy,” Nature, vol. 585, pp. 357–362, 2020. DOI: 10.1038/s41586-020-2649-2.
32. W. McKinney, “Data structures for statistical computing in Python,” in Proceedings of the 9th Python in Science Conference, pp. 56–61, 2010.
33. M. F. Akay, “Effective K-nearest neighbor classifications for Wisconsin breast cancer data sets,” Journal of the Chinese Institute of Engineers, vol. 43, no. 1, 2020. DOI: 10.1080/02533839.2019.1676658.
34. Sharma and D. Kumar, “Classification with 2-D convolutional neural networks for breast cancer diagnosis,” Scientific Reports, vol. 12, article 21857, 2022. DOI: 10.1038/s41598-022-26378-6.
35. S. Das and T. Roy Chowdhury, “HealthSync: A Web-Based Smart Clinical Appoint-ment and Lab Test Scheduling System,” International Journal of Emerging Trends in Engineering and Technology, vol. 15, pp. 888–893, 2026.
36. Q. Hu, H. M. Whitney, and M. L. Giger, “A deep learning methodology for im-proved breast cancer diagnosis using multiparametric MRI,” Scientific Reports, vol. 10, article 10536, 2020. DOI: 10.1038/s41598-020-67441-4.
37. R. El Shawi, K. Kilanava, and S. Sakr, “An interpretable semi-supervised framework for patch-based classification of breast cancer,” Scientific Reports, vol. 12, article 16734, 2022. DOI: 10.1038/s41598-022-20268-7.
38. J. Ferlay, M. Ervik, F. Lam, M. Laversanne, T. Colombet, M. Mery, M. Pin˜eros, A. Znaor, I. Soerjomataram, and F. Bray, “Global Cancer Observatory: Cancer Today,” International Agency for Research on Cancer, Lyon, France, 2024.
39. H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” CA: A Cancer Journal for Clinicians, vol. 71, no. 3, pp. 209–249, 2021. DOI: 10.3322/caac.21660.
40. S. Chakrabarti, T. Roy Chowdhury, P. Roy, and A. Nag, “Advanced DeepLung-CareNet: A Next-Generation Framework for Lung Cancer Prediction,” 2026. DOI: 10.1007/978-981-95-3671-9 9.