Breast Cancer Diagnosis Using Supervised Machine Learning for Benign and Malignant Classification

G. Naganandini
School of Computing and IT, REVA University, Bengaluru, Karnataka, India
R17pcs10@reva.edu.in (corresponding author)

Vishwanath R. Hulipalled
School of Computing and IT, REVA University, Bengaluru, Karnataka, India
vishwanth.rh@reva.edu.in

Received: 17 April 2025 | Revised: 10 May 2025 and 27 May 2025 | Accepted: 31 May 2025

 DOI: https://doi.org/10.48084/etasr.11563

ABSTRACT

This study investigates the application of supervised machine learning for classifying breast tumors as benign or malignant, leveraging the Breast Cancer Wisconsin Dataset. The proposed method encompasses a comprehensive pipeline, beginning with data preprocessing to address missing values and ensure feature normalization. Exploratory Data Analysis (EDA) techniques are employed to uncover patterns and relationships within the data. To enhance model performance, feature selection is performed using various techniques, including correlation-based selection, tree-based methods, and Recursive Feature Elimination with Cross-Validation (RFECV). Machine learning algorithms, including Random Forest (RF), SVM, Logistic Regression (LR), and Gradient Boosting (GB), were trained on the selected features. Hyperparameter tuning was performed using grid and randomized search to optimize model accuracy. The results demonstrate the effectiveness of the proposed method, achieving significant improvements in classification metrics such as precision, recall, F1-score, and ROC-AUC. These findings underscore the potential of machine learning to enhance diagnostic accuracy and reliability, offering a scalable, efficient, and robust approach to breast cancer diagnosis. This work paves the way for future integration of advanced techniques, including deep learning models and larger datasets, to further improve diagnostic outcomes and accessibility.

Keywords-breast cancer Wisconsin dataset; EDA; Recursive Feature Elimination with Cross-Validation (RFECV); precision; recall; F1-score; ROC-AUC

I. INTRODUCTION

Breast cancer is still one of the main causes of cancer-related diseases and death in women around the world [1]. Biopsy and mammography, two common ways to diagnose problems, are not always accurate, easy to access, or easy to understand. Machine Learning (ML) has made some exciting progress in finding complex data patterns that other methods might miss. This study investigates how to use supervised ML methods to classify tumors in the Breast Cancer Wisconsin Dataset as benign or malignant, focusing on data preprocessing, Exploratory Data Analysis (EDA), and feature selection methods. In particular, this study examines eight feature selection methods, investigating their particular contributions to classification accuracy. Cross-validation was used, running a complete ML pipeline comprising imputation, scaling, encoding, and hyperparameter tuning.

ML plays an increasingly important role in the diagnosis of breast cancer, especially in the classification of tumors as benign or malignant. Supervised ML models are particularly promising for improving diagnostic accuracy [1]. Although unsupervised methods such as clustering have been investigated [2], supervised ML continues to lead because of its effectiveness with labeled datasets. In [3], a comparative analysis confirmed that supervised models outperform semi-supervised ones, particularly on structured datasets such as the Breast Cancer Wisconsin Dataset. In addition, in [4], it was shown that combining feature extraction techniques with classification algorithms improves the diagnostic accuracy.

Several recent studies have focused on performance optimization. In [5], the success of models such as RF, LR, and Gradient Boosting (GB) in achieving high precision and F1-scores was highlighted. In [6], the importance of feature engineering and hyperparameter tuning in improving the results of the supervised model was underlined. The use of feature optimization techniques, such as Recursive Feature Elimination (RFE) and Chi-square testing, has been shown to improve results [7]. In [8], more than a 10% performance gain was

obtained from this preprocessing. Building effective ML pipelines with models such as SVM and Naïve Bayes (NB) also proves valuable [9]. At the same time, in [10], the limitations of conventional screening tools were emphasized, advocating for ML-based digital diagnostics.

In [11], ensemble methods were proposed as reliable solutions, while in [12, 13], hybrid and bioinspired models were explored to improve sensitivity and reduce false detections. In [14], the performance of LR paired with strong feature selection was reaffirmed. In [15], a deep learning framework was tailored for early-stage breast cancer detection. From a medical imaging standpoint, in [16], the use of ML to complement traditional radiological techniques was investigated. In [17], the use of explainable AI (XAI) was emphasized to increase trust and transparency in clinical applications. In [18], it was supported that modern algorithms enhance both prediction accuracy and understanding of feature importance. In [19], supervised models were benchmarked, providing foundational insights for future work.

In [20], ensemble learning was proposed to address class imbalance in datasets. In [21], a CNN-based metalearning approach was introduced to improve model generalizability and performance. Recent works have extended ML capabilities to areas such as spatio-temporal learning and advanced neural networks [22]. In [23], AST-GNN was developed for predictive modeling, while in [24], adaptive multiclass feature selection techniques were introduced. In [25, 26], novel deep learning and unsupervised feature selection approaches were proposed. Furthermore, in [27], enhancements were proposed for optimization algorithms, using a multipopulation whale optimization strategy. However, despite these technological advances, challenges remain around model generalizability and clinical integration.

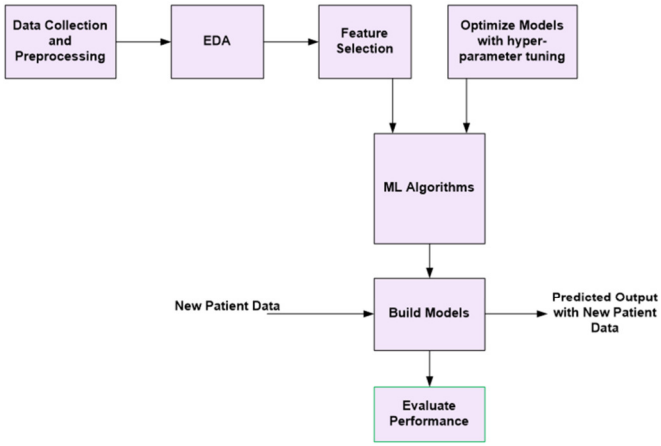

Fig. 1. Architecture of the proposed framework.

II. PROPOSED METHOD

This study employs a systematic process that involves data preparation, feature selection, and supervised ML model development. Figure 1 presents the method followed.

A. Data Collection and Preprocessing

The Breast Cancer Wisconsin Dataset is widely used for research and ML methods in breast cancer diagnosis. It includes aspirate sample characteristics and a target variable that indicates a benign or malignant diagnosis. For missing data, mean imputation was used, as:

$$Mean\ Imputation: x_i = \frac{\sum_{j=1}^{n} x_j}{n} \quad (1)$$

The features in a dataset may have varying ranges and magnitudes. As models such as LR or KNN are sensitive to such differences, min-max scaling was used to ensure fair contribution during model training.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2)$$

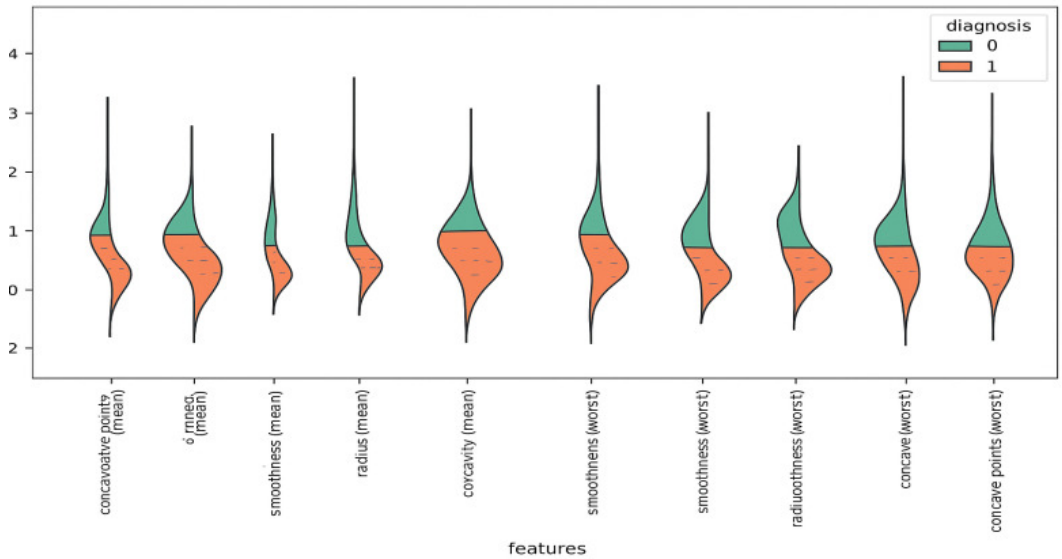

Fig. 2. Violin plot for distribution and density.

1) Encoding Categorical Variables

Since ML models cannot directly process non-numeric data, such as labels or categories, encoding converts them into numeric formats. One-hot encoding converts categories into binary vectors, making it suitable for non-ordinal data. For example, for a feature Color with categories {Red, Blue, Green}, one-hot encoding results in:

$$Red \rightarrow [1,0,0],\ Blue \rightarrow [0,1,0],\ Green \rightarrow [0,0,1]$$

Label encoding assigns a unique numeric value to each category, which is suitable for ordinal data. For example, for Education Level with categories {High School, Bachelor's, Master's}:

$$High\ School \rightarrow 1,\ Bachelor's \rightarrow 2,\ Master's \rightarrow 3$$

B. Exploratory Data Analysis (EDA)

This study highlights differences between benign and malignant tumor samples and examines model stability by using EDA techniques to find patterns and correlations in numerical data. A correlation matrix is computed, showing the correlation coefficients (r) for every pair of features:

$$Correlation\ r_{ij} = \frac{Cov(x_i,x_j)}{\sigma x_i.\sigma x_j} \quad (3)$$

where r_{ij} is the correlation between features x_i and x_j, Cov is the covariance, and σ is the standard deviation. The matrix is represented as a heatmap in Figure 3, with darker colors indicating stronger correlations.

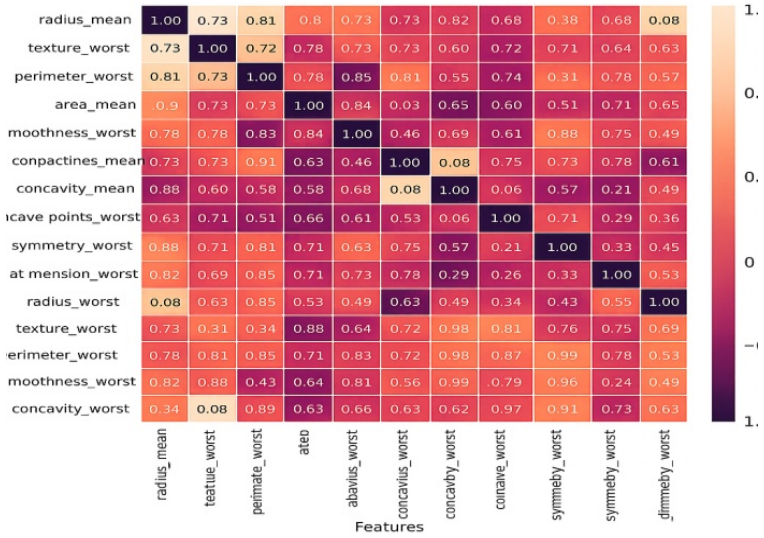

Fig. 3. Correlation heatmap.

C. Feature Selection

Feature selection is a vital step to identify and retain the most informative features from a dataset while removing redundant or irrelevant ones. Highly correlated features provide redundant information, and they were eliminated using:

$$|r_{ij}| > threshold \quad (4)$$

The Chi-Square statistic was computed for each feature to measure the dependency on the target variable.

$$x^2 = \sum \frac{(O-E)^2}{E} \quad (5)$$

where O is the observed frequency and E is the expected frequency. An RF was trained to calculate the feature importance I_j and rank features based on their scores.

$$I_j = \frac{1}{N}\sum_{t=1}^{N} \Delta G_t(j) \quad (6)$$

where $\Delta G_t(j)$ is the decrease in Gini impurity for feature j in tree t, and N is the total number of trees. A Lasso regression was used to penalize less important features and shrink their coefficients to zero.

$$\min\left(\frac{1}{2n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \lambda\sum_{j=1}^{p}|w_j|\right) \quad (7)$$

where w_j are feature weights, and λ is the regularization parameter. With their importance and contributions detailed in Table I and Figure 4, the vote-based choice found 16 essential features for classification, ensuring the model's accuracy and efficiency in detecting breast cancer.

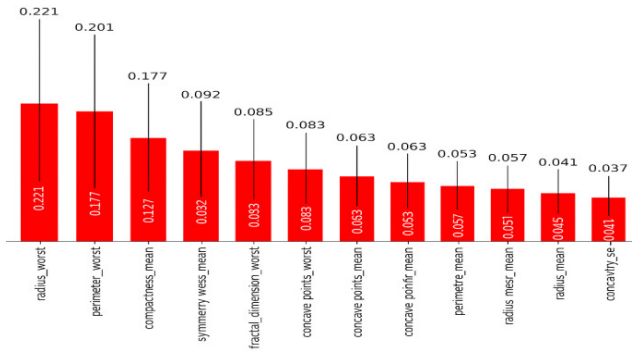

Fig. 4. Selection of the most important features.

TABLE I. VOTING-BASED SELECTED FEATURES

Votes	Index	Chi square	RF	Extratrees	RFE	RFECV	L1	Final R score
0	texture_mean	1	1	1	1	1	1	6
1	area_mean	1	1	1	1	1	1	6
7	area_se	1	1	1	1	1	1	6
3	concavity_mean	1	1	1	1	1	0	5
9	concavity_se	1	1	1	1	1	0	5
13	concavity_worst	1	1	1	1	1	0	5
14	symmetry_worst	1	1	1	1	1	0	5
12	smoothness_worst	1	1	1	1	0	0	4
15	fractal_dimension_worst	1	1	1	0	0	0	3
2	smoothness_mean	0	0	0	1	1	0	2
5	fractal_dimension_mean	0	0	1	0	1	0	2
4	symmetry_mean	1	0	0	0	0	0	1
8	smoothness_se	0	0	0	1	0	0	1
11	fractal_dimension_se	0	1	0	0	0	0	1
6	texture_se	0	0	0	0	0	0	0
10	symmetry_se	0	0	0	0	0	0	0

D. Hyperparameter Optimization

Hyperparameter tuning aims to increase model generalization and predictive accuracy. This study used grid search and randomized search.

E. Proposed Algorithm

1) Step 1: Data Collection and Preprocessing

- Input: Breast Cancer Wisconsin Data Set.
- Handle Missing Values: For each feature with missing values, calculate the mean or median using (1).
- Normalize Features: For each numerical feature x, apply min-max scaling using (2).
- Encode Categorical Variables:

  One-hot encoding for non-ordinal data:

  $If\ x \in \{a, b, c\},$ encode as $[1,0,0], [0,1,0], [0,0,1]$

  Label encoding for ordinal data:

  $x\ \in \{a, b, c\}, \rightarrow\ x' \in \{1,2,3\}$

2) Step 2: Exploratory Data Analysis (EDA)

- Visualize Distributions: Use Violin Plots to observe the density and spread for each feature.
- Correlation Analysis: Compute the correlation matrix using (3). Visualize the matrix using a heatmap to identify multicollinearity.

3) Step 3: Feature Selection

- Correlation-Based Selection: Identify feature pairs (X_i, X_j) using (3) and (4). Retain only one feature from highly correlated pairs.
- Chi-square Test: Compute the Chi-Square statistic for each feature with respect to the target (5).
- Recursive Feature Elimination (RFE): Fit a model $f(X)$, calculate feature importance scores, and iteratively remove the least important features until k features remain.
- RF-based Selection: Train an RF model to compute feature importance I_j (6) and retain features with $I_j > threshold$.
- L1 Regularization (Lasso): Fit a Lasso regression model using (7) and retain features with $w_j \neq 0$.

4) Step 4: Model Training and Evaluation

- Split Data: Divide the dataset into training and testing sets (e.g., 80:20 split).
- Train Models: Fit multiple machine learning models $f(X)$:

  $$f(X) = \hat{y} \quad (8)$$

  where $\hat{y} \epsilon \{Benign, Malignant\}$.
- Optimize Models: Use grid search or random search for hyperparameter tuning with cross-validation:

  $$cross - validation\ score = \frac{1}{k}\sum_{i=1}^{k} Accuracy_i \quad (9)$$

- Evaluation Metrics

  $$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (10)$$

  $$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (11)$$

  $$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (12)$$

  where TP, TN, FP, FN are True Positives, True Negatives, False Positives, and False Negatives.

  $$\text{F1} - \text{score} = 2 \times \frac{\text{Precision.Recall}}{\text{Precision+Recall}} \quad (13)$$

5) Step 5: Prediction

- Input: New patient data X_{new}.
- Preprocess Input: Apply the same preprocessing and feature selection steps as used in training.
- Make Predictions: Use the trained model $f(X)$ to predict:

  $$y_{new} = f(X_{new}) \quad (14)$$

- Output: Selected features, trained model, and predictions for new data y_{new}.

III. RESULTS AND DISCUSSION

The evaluation included 56 supervised ML models, with seven algorithms: LR, RF, SVM, Extra Trees (ET), KNN, DT, and NB. These models were applied across eight different feature selection strategies: Correlation, Chi-Square, Recursive Feature Elimination (RFE), Recursive Feature Elimination with Cross-Validation (RFECV), RF Importance, ET Importance, L1-based selection, and a voting-based method. The Breast Cancer Wisconsin dataset, which consists of 30 features and 569 patient records, was used for training and testing these models using stratified 10-fold cross-validation. Confusion matrices, precision-recall curves, and ROC-AUC were used to evaluate performance. With a ROC-AUC of 0.97 and an Average Precision (AP) of 0.99, the best-performing model was LR trained on features chosen using the RF importance method.

TABLE II. TRAINING PERFORMANCE - 5-FOLD CROSS-VALIDATION SCORES ON TRAINING DATA

Algorithm	Feature selection	Accuracy	Average precision	F1	Precision	Recall	ROC-AUC
LR	Correlation	0.9775	0.9936	0.9700	0.9804	0.9611	0.9939
SVM	Chi2	0.9648	0.9914	0.9532	0.9670	0.9411	0.9936
SVM	RFE	0.9626	0.9908	0.9481	0.9739	0.9286	0.9930
LR	RFECV	0.9724	0.9933	0.9632	0.9800	0.9477	0.9938
LR	RF	0.9774	0.9945	0.9704	0.9742	0.9675	0.9947
LR	Extra trees	0.9725	0.9931	0.9635	0.9752	0.9546	0.9931
LR	L1	0.8994	0.9514	0.8601	0.9214	0.8088	0.9585
LR	Voted	0.9725	0.9932	0.9635	0.9752	0.9546	0.9935

TABLE III. TESTING PERFORMANCE

Algorithm	Feature Selection	Accuracy	F1	Precision	Recall	ROC-AUC
LR	Correlation	0.9591	0.9412	0.9492	0.9333	0.9532
SVM	Chi2	0.9591	0.9421	0.9344	0.9500	0.9570
SVM	RFE	0.9649	0.9492	0.9655	0.9333	0.9577
LR	RFECV	0.9591	0.9402	0.9649	0.9167	0.9493
LR	RF	0.9766	0.9661	0.9828	0.9500	0.9705
LR	Extra trees	0.9649	0.9492	0.9655	0.9333	0.9577
LR	L1	0.9064	0.8621	0.8929	0.8333	0.8896
LR	Voted	0.9591	0.9412	0.9492	0.9333	0.9532

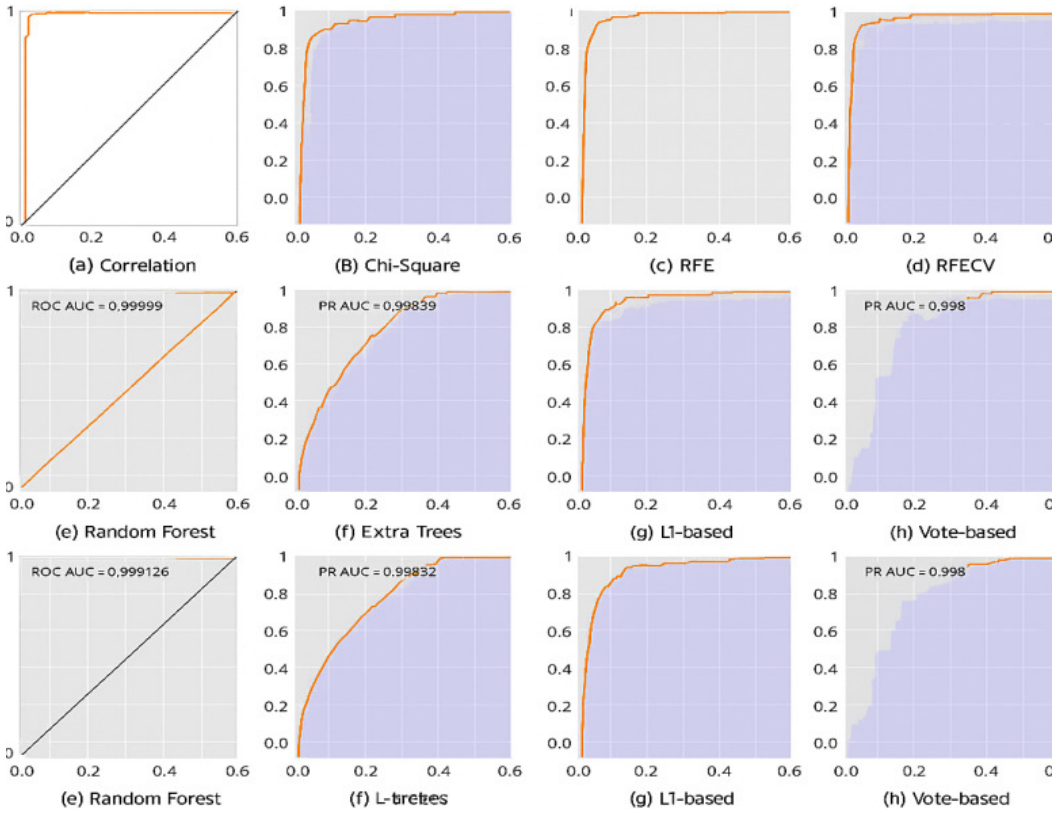

Fig. 5. Feature selection on various methods.

Figures 6-8 demonstrate the performance of LR as the best model when combined with RF feature selection. This linear model removes unnecessary variables and works well for binary classification tasks. According to the AUC graph comparison, LR has the highest AUC, demonstrating its capacity to efficiently utilize specific features. The graph highlights how crucial it is to choose the best feature selection method and combine it with an ML algorithm that works well with it. The results of LR are likely because of its simplicity and interpretability, which make it suitable for binary classification tasks, feature selection techniques that reduce dimensionality and noise, thus supporting linear models. In addition, LR is difficult to overfit on small datasets such as the Breast Cancer Wisconsin dataset. On the other hand, L1 regularization performed poorly, possibly because of over-penalization and feature removal. SVM and other models are sensitive to feature scaling and may need more adjustment.

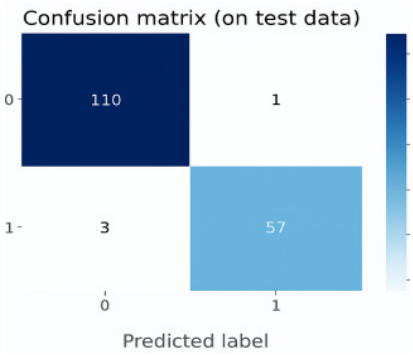

Fig. 6. Confusion matrix.

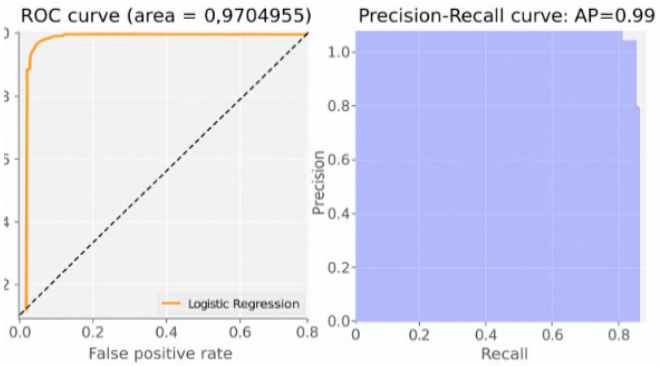

Fig. 7. Best model: LR with RF feature selection.

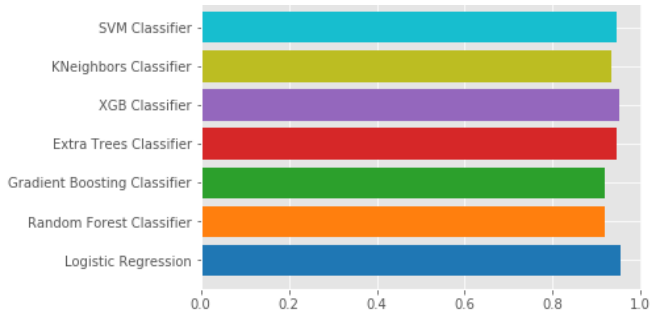

Fig. 8. AUC comparison of ML algorithms with RF feature selection.

IV. CONCLUSION

Using supervised ML techniques, this study offers a rigorous framework for the diagnosis of breast cancer. Using the Breast Cancer Wisconsin Dataset, thorough preprocessing techniques were employed to handle missing data, standardize features, and encode categorical variables. Insights into the dataset were obtained by EDA. Feature selection was used to keep only the most predictive features. Model performance was enhanced by hyperparameter tuning using grid search and randomized search. Using 16 features, many models were trained, optimized, and evaluated. With a test data accuracy of 0.977 and an AUC of 0.971, LR produced the best results. This approach shows how well combining data preprocessing, visualization, and ML methods can achieve precise breast cancer detection. The detailed pipeline for breast cancer classification using supervised ML methods shows how meticulous technique integration can produce very accurate results. However, the model has drawbacks, including bias in the dataset, generalizability, clinical interpretability, and ethical issues. Future studies should integrate electronic health record systems, incorporate real-time patient data, and externally validate multiple datasets.

V. ETHICAL CONSIDERATIONS AND DATA PRIVACY

Using medical datasets such as the Breast Cancer Wisconsin Dataset calls for close consideration of patient privacy and ethical data use. Although the dataset is public and anonymized, any use of diagnostic models in actual situations has to follow healthcare laws, such as HIPAA (Health Insurance Portability and Accountability Act) or GDPR, therefore ensuring data protection and informed consent. Avoiding biases that could harm certain patient populations and considering algorithmic fairness are also very important.

REFERENCES

[1] C. G. Yedjou, S. S. Tchounwou, R. A. Aló, R. Elhag, B. Mochona, and L. Latinwo, "Application of Machine Learning Algorithms in Breast Cancer Diagnosis and Classification," *International Journal of Science Academic Research*, vol. 2, no. 1, pp. 3081–3086, Jan. 2021.

[2] W. C. Shia, L. S. Lin, and D. R. Chen, "Classification of malignant tumours in breast ultrasound using unsupervised machine learning approaches," *Scientific Reports*, vol. 11, no. 1, Jan. 2021, Art. no. 1418, https://doi.org/10.1038/s41598-021-81008-x.

[3] N. Al-Azzam and I. Shatnawi, "Comparing supervised and semi-supervised Machine Learning Models on Diagnosing Breast Cancer," *Annals of Medicine and Surgery*, vol. 62, pp. 53–64, Feb. 2021, https://doi.org/10.1016/j.amsu.2020.12.043.

[4] M. S. Gharibdousti, S. M. Haider, D. Ouedraogo, and S. Lu, "Breast cancer diagnosis using feature extraction techniques with supervised and unsupervised classification algorithms," *Applied Medical Informatics*, vol. 41, no. 1, pp. 40–52, Mar. 2019.

[5] Z. Q. Fanoos and J. Abdulhadi, "Breast cancer diagnosis using supervised machine learning classification algorithms," presented at the 2nd International Conference on Applied Research and Engineering (ICARAE2022), Cape Town, South Africa, 2023, https://doi.org/10.1063/5.0167964.

[6] P. Dadheech, V. Kalmani, S. R. Dogiwal, V. K. Sharma, A. Kumar, and S. K. Pandey, "Breast cancer prediction using supervised machine learning techniques," *Journal of Information and Optimization Sciences*, vol. 44, no. 3, pp. 383–392, 2023, https://doi.org/10.47974/JIOS-1348.

[7] K. M. M. Uddin, N. Biswas, S. T. Rikta, and S. K. Dey, "Machine learning-based diagnosis of breast cancer utilizing feature optimization technique," *Computer Methods and Programs in Biomedicine Update*,

vol. 3, 2023, Art. no. 100098, https://doi.org/10.1016/j.cmpbup.2023.100098.

[8] V. Chaurasia and S. Pal, "Applications of Machine Learning Techniques to Predict Diagnostic Breast Cancer," *SN Computer Science*, vol. 1, no. 5, Sep. 2020, Art. no. 270, https://doi.org/10.1007/s42979-020-00296-8.

[9] O. M. William Wolberg, "Breast Cancer Wisconsin (Diagnostic)." UCI Machine Learning Repository, 1993, https://doi.org/10.24432/C5DW2B.

[10] O. Ginsburg *et al.*, "Breast cancer early detection: A phased approach to implementation," *Cancer*, vol. 126, no. S10, pp. 2379–2393, May 2020, https://doi.org/10.1002/cncr.32887.

[11] A. E. Kilic and M. Karakoyun, "Breast Cancer Detection Using Machine Learning Algorithms," presented at the 2nd International Conference on Scientific and Academic Research, Konya, Turkey, Mar. 2023.

[12] S. Devi, R. Kaul Ghanekar, J. Pande, D. Dumbre, R. Chavan, and H. Gupta, "Prediction and Diagnosis of Breast Cancer Using Machine and Modern Deep Learning Models," *Asian Pacific Journal of Cancer Prevention*, vol. 25, no. 3, pp. 1077–1085, Mar. 2024, https://doi.org/10.31557/APJCP.2024.25.3.1077.

[13] R. S. P. Priya and P. S. Vadivu, "Bio-inspired ensemble feature selection (biefs) and kernel extreme learning machine classifier for breast cancer diagnosis," *International Journal of Health Sciences*, pp. 1404–1429, Jun. 2022, https://doi.org/10.53730/ijhs.v6nS5.8976.

[14] S. Sakib, N. Yasmin, A. K. Tanzeem, F. Shorna, K. Md. Hasib, and S. B. Alam, "Breast Cancer Detection and Classification: A Comparative Analysis Using Machine Learning Algorithms," in *Proceedings of Third International Conference on Communication, Computing and Electronics Systems*, vol. 844, V. Bindhu, J. M. R. S. Tavares, and K. L. Du, Eds. Springer Singapore, 2022, pp. 703–717.

[15] A. Bekkouche, M. Merzoug, M. Hadjila, and W. Ferhi, "Towards Early Breast Cancer Detection: A Deep Learning Approach," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 17517–17523, Oct. 2024, https://doi.org/10.48084/etasr.8634.

[16] S. Iranmakani *et al.*, "A review of various modalities in breast imaging: technical aspects and clinical outcomes," *Egyptian Journal of Radiology and Nuclear Medicine*, vol. 51, no. 1, Dec. 2020, Art. no. 57, https://doi.org/10.1186/s43055-020-00175-5.

[17] T. Islam *et al.*, "Predictive modeling for breast cancer classification in the context of Bangladeshi patients by use of machine learning approach with explainable AI," *Scientific Reports*, vol. 14, no. 1, Apr. 2024, Art. no. 8487, https://doi.org/10.1038/s41598-024-57740-5.

[18] Y. Guo *et al.*, "Machine learning and new insights for breast cancer diagnosis," *Journal of International Medical Research*, vol. 52, no. 4, Apr. 2024, Art. no. 03000605241237867, https://doi.org/10.1177/03000605241237867.

[19] M. Gupta and B. Gupta, "A Comparative Study of Breast Cancer Diagnosis Using Supervised Machine Learning Techniques," in *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, Feb. 2018, pp. 997–1002, https://doi.org/10.1109/ICCMC.2018.8487537.

[20] R. Rabiei, "Prediction of Breast Cancer using Machine Learning Approaches," *Journal of Biomedical Physics and Engineering*, vol. 12, no. 3, Jul. 2022, https://doi.org/10.31661/jbpe.v0i0.2109-1403.

[21] M. D. Ali *et al.*, "Breast Cancer Classification through Meta-Learning Ensemble Technique Using Convolution Neural Networks," *Diagnostics*, vol. 13, no. 13, Jun. 2023, Art. no. 2242, https://doi.org/10.3390/diagnostics13132242.

[22] T. R. Mahesh *et al.*, "Transformative Breast Cancer Diagnosis using CNNs with Optimized ReduceLROnPlateau and Early Stopping Enhancements," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, Jan. 2024, Art. no. 14, https://doi.org/10.1007/s44196-023-00397-1.

[23] H. Zhou, D. Ren, H. Xia, M. Fan, X. Yang, and H. Huang, "AST-GNN: An attention-based spatio-temporal graph neural network for Interaction-aware pedestrian trajectory prediction," *Neurocomputing*, vol. 445, pp. 298–308, Jul. 2021, https://doi.org/10.1016/j.neucom.2021.03.024.

[24] M. Fan, X. Zhang, J. Hu, N. Gu, and D. Tao, "Adaptive Data Structure Regularized Multiclass Discriminative Feature Selection," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 10, pp. 5859–5872, Jul. 2022, https://doi.org/10.1109/TNNLS.2021.3071603.

[25] D. Zou *et al.*, "Deep Field Relation Neural Network for click-through rate prediction," *Information Sciences*, vol. 577, pp. 128–139, Oct. 2021, https://doi.org/10.1016/j.ins.2021.06.079.

[26] P. Zhou, J. Chen, M. Fan, L. Du, Y. D. Shen, and X. Li, "Unsupervised feature selection for balanced clustering," *Knowledge-Based Systems*, vol. 193, Apr. 2020, Art. no. 105417, https://doi.org/10.1016/j.knosys.2019.105417.

[27] M. A. Navarro *et al.*, "An improved multi-population whale optimization algorithm," *International Journal of Machine Learning and Cybernetics*, vol. 13, no. 9, pp. 2447–2478, Sep. 2022, https://doi.org/10.1007/s13042-022-01537-3.