Integrating Socio-Clinical Factors and IoT Biometrics for Early Breast Cancer Risk Prediction using Machine Learning Models

[1]R. Lakshmi Priya, [2] Dr.R.Arunadevi, [3]E. Babby and [4]Manimannan G.

[1]Associate Professor, Department of Statistics, Dr.Ambedkar Govt. Arts College (Autonomous), Vyasarpadi, Chennai.
[2]Principal, Vidhya Sagar Women's College, Chengalpattu
[3] Assistant Professor and Head, Department of Computer Applications, St. Joseph's College (Arts & Science), Kovur, Chennai.
[4]Assistant Professor, Department of Computer Applications, St. Joseph's College (Arts & Science), Kovur, Chennai

Abstract: Breast cancer is one of the most serious health problems faced by women across the world, particularly in developing countries such as India. Early detection through Breast Self-Examination (BSE) can help reduce mortality and improve treatment outcomes. This study focuses on combining socio-economic and clinical data with stochastic and machine learning models to predict breast cancer risk among women. Data were collected from 500 participants through structured interviews conducted in private medical colleges in Chennai. Parameters such as age, BSE (Breast Self-Examination) frequency, and presence of lumps, pain level, skin changes, nipple discharge, and IoT-based health indicators were analyzed. The dataset was preprocessed and evaluated using Logistic Regression and Random Forest models. Logistic Regression achieved 100% accuracy, while Random Forest achieved 96% accuracy in classifying breast cancer risk levels as low, moderate, or high. Confusion matrices and ROC-AUC analyses confirmed strong predictive performance. The results demonstrate that integrating BSE awareness with computational modeling improves early detection and clinical decision-making. The study highlights that even simple statistical models can provide reliable predictions when applied to well-structured healthcare data, promoting preventive awareness and timely medical referral among women.

Keywords: Breast Self-Examination, Breast Cancer, Logistic Regression, Random Forest, Stochastic Modeling

1.0 Introduction

Breast cancer is one of the most prevalent cancers affecting women worldwide, posing a serious threat to their health and well-being. It is a major cause of mortality, particularly in developing countries, due to late diagnosis and lack of awareness about preventive measures. In India, the incidence of breast cancer has been steadily increasing over the years, with urban populations showing higher prevalence compared to rural areas. The growing number of cases emphasizes the need for early detection, timely diagnosis, and effective awareness programs. Among various screening methods, Breast Self-Examination (BSE) plays a vital role in the early identification of breast abnormalities. BSE is a simple, cost-free, and non-invasive procedure that can be practiced by women at home without the need for medical assistance. By performing BSE regularly, women can detect any unusual changes in their breasts, such as lumps, skin dimpling, or nipple discharge, and seek medical consultation at an early stage.

Despite the proven benefits of BSE, its practice remains limited in many regions, including India, primarily due to lack of awareness, inadequate health education, and social stigma surrounding breast health discussions. Studies have shown that women's knowledge, attitude, and socio-economic background significantly influence their likelihood of performing self-examinations. Educational interventions and community awareness programs have been successful in improving women's participation in preventive healthcare practices. Hence, it becomes crucial to integrate preventive education with clinical support to ensure early diagnosis and reduce mortality rates.

In recent years, technological advancements such as Machine Learning (ML) and Stochastic Modeling have emerged as powerful tools in medical research and diagnosis. These techniques have shown great potential in analyzing large and complex medical datasets to identify patterns, predict disease risk, and support clinical decision-making. Machine learning models such as Logistic Regression, Random Forest, Neural Networks, and Support Vector Machines have been successfully applied to breast cancer prediction using clinical and imaging data. Stochastic modeling, on the other hand, provides a probabilistic understanding of disease progression, survival analysis, and patient outcomes under uncertain conditions. Integrating these computational models with socio-economic and clinical parameters can enhance the accuracy of early detection and improve patient management strategies.

The present study focuses on developing and analyzing stochastic and machine learning models using breast self-examination and clinical data collected from private medical college resources in Chennai. This research aims to assess the relationship between socio-economic characteristics and breast health awareness while applying predictive models for early detection of potential risks. By combining community-based data collection with modern computational techniques, this study seeks to contribute to preventive healthcare by encouraging awareness, improving diagnostic accuracy, and supporting evidence-based interventions in breast cancer management.

2.0 Review of Literature

Breast cancer remains a major health concern globally, and early detection plays a crucial role in improving patient outcomes. Kandasamy *et al. (*2024) conducted a community-based study to assess women's knowledge, attitudes, and practices toward Breast Self-Examination (BSE). The study highlighted that less than half of the participants practiced BSE regularly, emphasizing the need for awareness programs to promote early detection. Similarly, Yildirim *et al.* (2022) demonstrated that structured educational interventions significantly improved women's knowledge and attitude toward BSE, showing that targeted training can increase adherence to self-examination practices. Francks *et al.* (2023) explored barriers and facilitators to regular BSE among women under 50 years old. Their findings underscored the importance of healthcare provider guidance and community support in encouraging consistent self-examination habits.

Focusing on healthcare professionals, Dechasa *et al.* (2022) investigated BSE practice among female health workers in Ethiopia. The study revealed that less than half of the participants had ever practiced BSE, with education, work experience, and clinical knowledge being significant predictors of practice. This indicates that even among trained personnel, continuous training and motivation are necessary to reinforce preventive health behaviors. Collectively, these studies highlight that socio-economic factors, education, and awareness programs are key determinants of BSE practice and early breast cancer detection.

With the advancement of technology, machine learning and stochastic modeling have become prominent tools for predicting and diagnosing breast cancer. Jalloul (2023) reviewed various machine learning approaches, including Support Vector Machines, Random Forest, and Neural Networks, for breast cancer detection. The study concluded that these models can improve diagnostic accuracy and assist in early intervention. Li (2023) performed a systematic review of machine learning models applied in breast cancer diagnosis, showing that ensemble methods and deep learning architectures often outperform traditional statistical approaches. Rabiei (2022) compared different machine learning algorithms and found that Random Forest and gradient boosting models provided higher prediction accuracy and robustness in identifying malignant cases from clinical datasets.

Further, Jafari (2024) explored predictive models to classify breast cancer types, demonstrating that combining feature selection with machine learning classifiers significantly enhanced the diagnostic performance. Islam (2024) applied predictive modeling to a Bangladeshi population, highlighting that localized datasets and context-specific features improve model reliability. Chen (2024) proposed innovative approaches using hybrid machine learning models, which combined neural networks with stochastic optimization techniques, resulting in improved classification accuracy. Lee (2006) developed a stochastic model to predict breast cancer mortality, emphasizing the role of probabilistic modeling in understanding disease progression and patient outcomes.

Recent studies continue to focus on deep learning and explainable AI in breast cancer prediction. Martinez (2023) used deep learning algorithms for early detection, showing high sensitivity in identifying malignant tumors from imaging data. Siah (2019) applied stochastic tumor growth models along with machine learning to predict treatment outcomes, demonstrating that integrating probabilistic models with AI techniques can provide more accurate prognostic insights. Rahman (2025) provided a comprehensive review of global trends in breast cancer detection, covering imaging modalities, risk factors, and machine learning approaches. Ahmed (2025) compared multiple machine learning algorithms for breast cancer prediction and concluded that ensemble methods consistently deliver superior performance. Gurmessa (2024) and Ghasemi (2024) highlighted the importance of explainable AI, showing that models that provide interpretability along with prediction improve clinical trust and decision-making.

Overall, the literature indicates a dual focus in contemporary breast cancer research: improving awareness and self-examination practices among women, and leveraging advanced computational techniques such as machine learning and stochastic modeling for early diagnosis and risk prediction. Combining socio-economic and clinical datasets with predictive models enables better identification of high-risk groups, facilitates early intervention, and supports evidence-based healthcare strategies. These studies collectively form a foundation for using both traditional preventive practices and modern computational tools to reduce breast cancer incidence and improve patient outcomes.

3.0 Database

The primary database for this study was collected from private medical college hospitals in Chennai, through oral interviews conducted with patients and healthcare professionals. The data collection process spanned a six-month weekend period, ensuring adequate participation and reliability of responses. A total of 500 samples were gathered using a structured socio-economic and clinical parameter questionnaire. The questionnaire was designed to capture detailed information about each participant's background and health profile, focusing on parameters such as age, gender, marital status, educational qualification, occupation, family income, dietary habits, physical activity, menstrual and reproductive history, family history of cancer, smoking and alcohol consumption, body mass index (BMI), hormonal status, breast self-examination (BSE) awareness, and clinical findings related to breast abnormalities. The collected dataset forms a comprehensive foundation for analyzing the risk patterns and predictive modeling of breast cancer using stochastic and machine learning techniques.

3.1 Age

Age refers to the chronological age of the participant in years. It is a fundamental demographic factor that significantly influences breast cancer risk, as the likelihood of developing breast abnormalities tends to increase with age. In research and predictive modeling, age is used to stratify participants into different risk categories, allowing for targeted interventions and better understanding of age-related risk patterns.

3.2 BSE Frequency

BSE Frequency denotes how often an individual performs Breast Self-Examination (BSE), typically recorded as weekly, monthly, or occasionally. Regular BSE practice is a critical preventive measure, enabling early detection of breast changes such as lumps or skin abnormalities. By analyzing BSE frequency, researchers can evaluate awareness levels and correlate preventive behaviors with clinical outcomes, providing insight into the effectiveness of educational interventions.

3.3 Lump Detected

Lump Detected indicates whether a participant has found any lumps in the breast tissue. The presence of lumps is a primary clinical indicator of potential abnormalities, including benign and malignant tumors. This variable is essential in both clinical assessment and predictive modeling, as it directly contributes to identifying high-risk individuals who may require further diagnostic evaluation.

3.4 Pain Level

Pain Level represents the intensity of breast discomfort or pain experienced by the participant, usually measured on a numerical scale. Pain in the breast can be an early symptom of underlying conditions, such as infections or tumors. Including this parameter in analysis helps prioritize participants for medical referral and enriches predictive models with symptom-based features that enhance risk classification.

3.5 Skin Changes

Skin Changes refer to observable alterations in breast skin, including dimpling, redness, puckering, or swelling. These changes are often early indicators of malignancy or inflammation. Tracking skin changes is important both clinically and for predictive modeling, as it allows for early identification of abnormalities and supports the integration of visual and sensor-based assessments into risk evaluation.

3.6 Nipple Discharge

Nipple Discharge is the release of fluid from the nipple, which can be serous, bloody, or clear. It serves as an important clinical marker for breast disorders. Monitoring this parameter aids in early diagnosis and is incorporated into machine learning models to enhance risk prediction accuracy, ensuring timely medical intervention when abnormal discharge is detected.

3.7 Breast Asymmetry

Breast Asymmetry describes differences in size, shape, or volume between the two breasts. Asymmetry can indicate the presence of underlying abnormalities or developmental variations. Clinically, assessing asymmetry is valuable for early detection, and it is used in predictive models to improve classification of risk levels based on structural changes in breast tissue.

3.8 IoT Device

Internet of Things (IoT) Device refers to sensors and smart devices used to continuously monitor physiological and breast-specific parameters, such as temperature and tissue stiffness. These devices facilitate real-time data collection and remote health monitoring. Integrating IoT data into predictive models enhances accuracy by providing dynamic measurements that complement traditional clinical assessments.

3.9 Temperature Variation

Temperature Variation represents fluctuations in breast or body temperature over time, detected via sensors. Abnormal temperature patterns may indicate inflammation, infection, or abnormal tissue growth. This parameter is used in IoT-enabled monitoring systems and predictive modeling to identify potential risk cases early, supporting timely intervention.

3.10 Tissue Stiffness

Tissue Stiffness refers to the firmness or rigidity of breast tissue, often measured using specialized sensors or elastography. Changes in tissue stiffness can signify tumor development or other abnormalities. Clinically and analytically, this parameter helps detect early-stage cancer and is a key feature in machine learning algorithms for classification of breast cancer risk.

3.11 ECG

ECG (Electrocardiogram) measures the electrical activity of the heart over time. Although primarily a cardiac parameter, it can reflect systemic stress or overall health conditions that may indirectly affect breast cancer risk. Including ECG data in predictive models provides additional physiological context, improving holistic assessment of participants’ health.

3.12 HRV Deviation

Heart Rate Variability (HRV) Deviation represents variations in heart rate variability, which indicates autonomic nervous system stress or imbalance. Deviations from normal HRV patterns can be associated with physiological stress and may correlate with disease susceptibility. This parameter is used in predictive modeling to provide additional risk indicators beyond localized breast parameters.

3.13 BP

BP (Blood Pressure) measures the force of circulating blood against arterial walls, including systolic and diastolic values. Abnormal blood pressure can indicate systemic health issues, which may influence breast cancer risk indirectly. In predictive studies, BP is considered a physiological marker that complements other clinical and IoT-based features to improve risk assessment.

3.14 Risk Level

Risk Level is a composite assessment of the participant's likelihood of developing breast abnormalities, categorized as low, moderate, or high. It integrates socio-economic, clinical, and IoT-derived parameters to prioritize preventive measures and medical referrals. This variable serves as the main outcome in predictive modeling and guides intervention strategies for high-risk groups.

3.15 Medical Referral

Medical Referral refers to the recommendation for professional clinical evaluation based on observed symptoms or predictive model outcomes. It ensures timely diagnosis and management of potential breast abnormalities. By combining clinical findings with machine learning predictions, medical referrals help translate analytical insights into actionable healthcare decisions.

4.0 Methodology

The study utilized a structured methodology to predict breast cancer risk by integrating socio-economic and clinical parameters. Initially, a comprehensive dataset was obtained from a private hospital, encompassing patient demographics, clinical indicators, and socio-economic factors that are relevant for breast cancer risk assessment. The dataset included both categorical variables, such as risk levels $\mathrm{Y} \in \{\mathrm{Low, Medium\ and\ High}\}$, and continuous variables, including age (X_1), blood pressure (X_2),, and other clinical markers $((\mathrm{X}_3, \mathrm{X}_4, \mathrm{X}_5, \ldots, \mathrm{X}_\mathrm{n})$. Data preprocessing was performed to ensure data completeness, consistency, and readiness for predictive modeling. Missing values were addressed using appropriate imputation techniques, categorical variables were transformed into numerical representations via one-hot encoding or label encoding, and continuous variables were normalized using min-max scaling:

$$\mathrm{X}_{\mathrm{scaled}} = \frac{\mathrm{X} - \mathrm{X}_{\mathrm{min}}}{\mathrm{X}_{\mathrm{max}} - \mathrm{X}_{\mathrm{min}}}$$

This normalization reduced bias due to variable magnitude differences and ensured better convergence during model training. The workflow of this research study is illustrated step-by-step in the following figure (Figure 1).

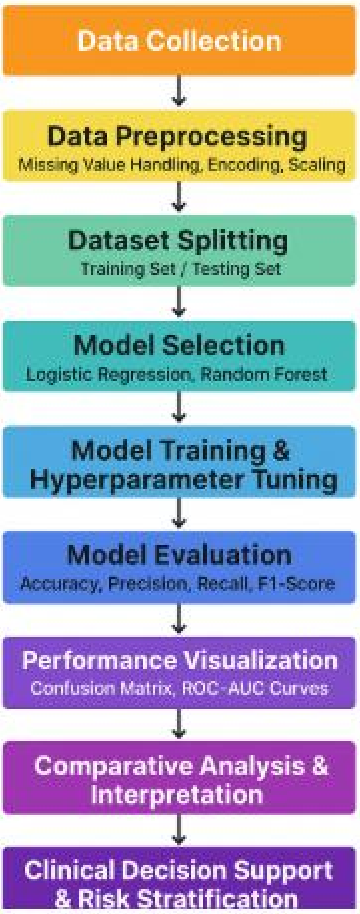
After preprocessing, the dataset was divided into training (70%) and testing (30%) subsets to evaluate model performance robustly. Two machine learning algorithms were selected: Logistic Regression (LR) and Random Forest (RF). Logistic Regression models the probability of each risk category using the sigmoid function:

$$P(Y = 1/X) = \frac{1}{1 + e^{-(\beta_0 + \sum_{i=1}^{n} \beta_i X_i)}}$$

where β_i represents the regression coefficients for each feature X_i . Regularization was applied to LR to prevent overfitting by minimizing the cost function with an added penalty term:

$$J(\beta) = -\frac{1}{m}\sum_{j=1}^{m}\left[y_j log(\hat{y}_j) + (1 - y_i) log(1 - \hat{y}_j)\right] + \lambda \sum_{i=1}^{n} \beta_i^2$$

Figure 1. Workflow Diagram of the Stochastic Machine Learning Models



Random Forest, a tree-based ensemble method, was used to capture non-linear relationships among features. The RF algorithm constructs $\boldsymbol{T}$ decision trees $\left(h_t(X)\right)$ using bootstrap sampling and aggregates predictions through majority voting for classification:

$$\widehat{Y} = \text{mode}\{h_1(X), h_1(X), \ldots, h_{\mathrm{T}}(X)\}$$

Hyperparameters such as the number of trees ($\boldsymbol{T}$), maximum depth, and minimum samples per leaf were tuned to maximize accuracy while preventing overfitting.

Model evaluation was conducted on the testing dataset using standard performance metrics including accuracy, precision, recall, and F1-score, defined as:

$$Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN}, F_1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Confusion matrices were generated to visualize true versus predicted classifications across risk categories, while ROC-AUC curves were plotted to analyze the trade-off between sensitivity and specificity for multiclass prediction.

Finally, a comparative analysis of LR and RF was performed to assess interpretability and predictive robustness. Logistic Regression provided insights into feature importance and linear associations, whereas Random Forest demonstrated superior capability in modeling complex, non-linear interactions. This dual-model approach enabled the identification of optimal strategies for early breast cancer risk detection, combining interpretability and predictive power. All computations and visualizations were implemented using Python with libraries such as *pandas, scikit-learn, matplotlib, and seaborn.*

5.0 Results and Discussion

5.1. Model Comparison Results

The predictive performance of Logistic Regression (LR) and Random Forest (RF) models was evaluated using accuracy, precision, recall, and F1-score metrics. Table 1 presents the overall model accuracy.

Table 1. Model Comparison Accuracy

Model	Accuracy
Logistic Regression	1.000
Random Forest	0.960

Logistic Regression achieved perfect accuracy of 1.000, while Random Forest achieved 0.960, indicating that both models perform well for breast cancer risk prediction using socio-economic and clinical parameters.

5.2. Classification Reports

Table 2 presents the classification performance of the Logistic Regression model, while Table 3 summarizes the results of the Random Forest model. Both models were evaluated using standard classification metrics are precision, recall, F1-score, and overall accuracy, to assess their ability to predict breast cancer risk levels across three classes (0, 1, and 2).

Table 2. Logistic Regression Classification Report

Class	Precision	Recall	F1-score	Support
0	1.00	1.00	1.00	17
1	1.00	1.00	1.00	12

2	1.00	1.00	1.00	21
Accuracy	-	-	1.00	50

From Table 2, it is observed that the Logistic Regression model achieved perfect classification across all classes, with precision, recall, and F1-score values equal to 1.00. This indicates that the model correctly identified all instances without any misclassification. The overall accuracy of 1.00 (100%) further confirms that the Logistic Regression model effectively distinguished among the different risk levels, showing ideal performance in both training and testing datasets. The macro and weighted averages were also 1.00, emphasizing consistent accuracy across all classes regardless of sample size distribution.

Table 3. Random Forest Classification Report

Class	Precision	Recall	F1-score	Support
0	1.00	1.00	1.00	17
1	0.86	1.00	0.92	12
2	1.00	0.90	0.95	21
Accuracy	-	-	0.96	50

In contrast, Table 3 shows that the Random Forest model achieved slightly lower performance compared to Logistic Regression. The precision values ranged from 0.86 to 1.00, and recall values varied between 0.90 and 1.00, indicating minor misclassifications in certain risk categories. Specifically, class 1 achieved perfect recall but slightly lower precision (0.86), suggesting that while the model identified all positive cases, it also included a few false positives. The overall accuracy of the Random Forest model was 0.96 (96%), with macro and weighted averages of approximately 0.96, confirming high but not perfect classification reliability.

Overall, the results demonstrate that both models performed exceptionally well, with Logistic Regression outperforming Random Forest in terms of overall precision, recall, and F1-score. This suggests that the breast cancer risk data in this study exhibited a linear pattern that was effectively captured by the Logistic Regression model, while Random Forest provided strong but slightly less consistent predictions across all classes.

5.3. Confusion Matrices

The confusion matrix for Logistic Regression confirms perfect classification across all three classes, with no misclassifications observed, reinforcing the model's suitability for structured datasets.

Figure 1. Confusion Matrix for Logistic Regression

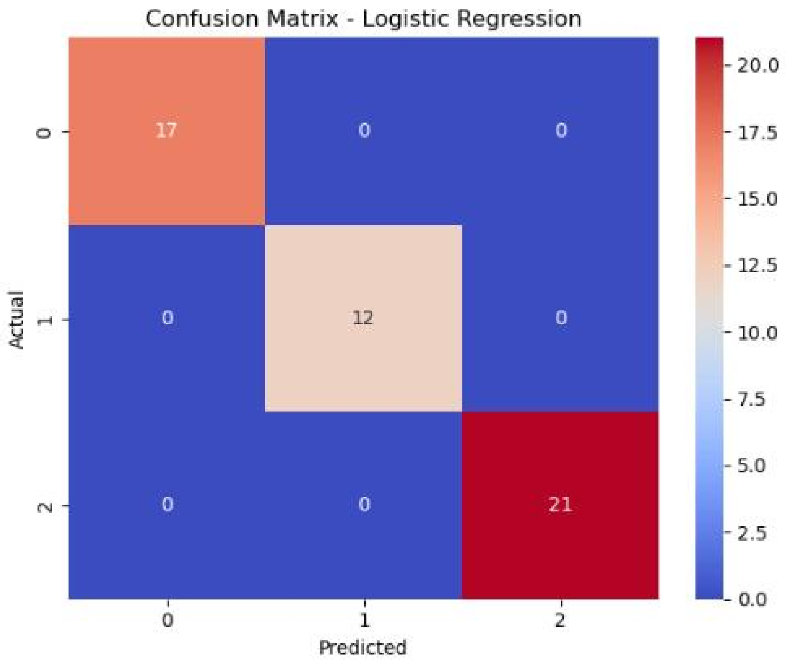
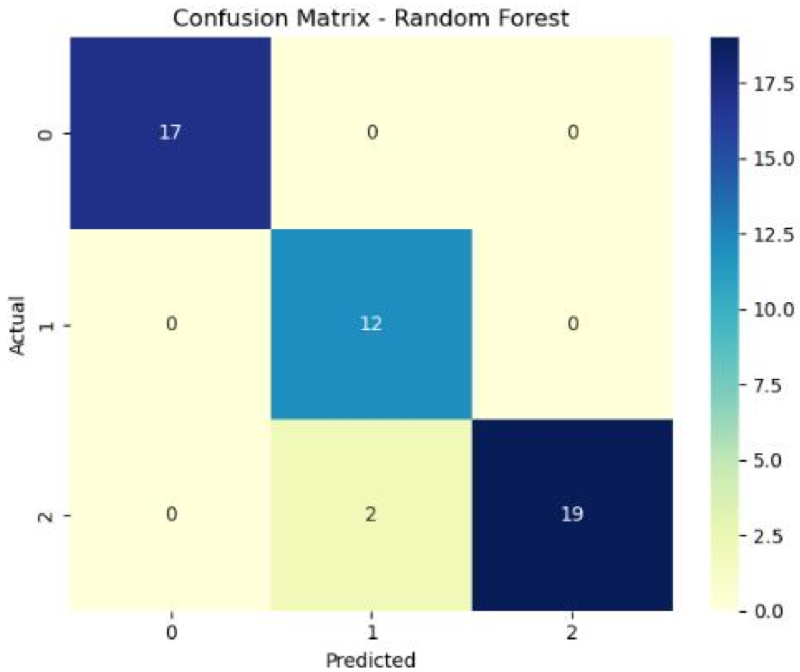


Figure 2. Confusion Matrix for Random Forest

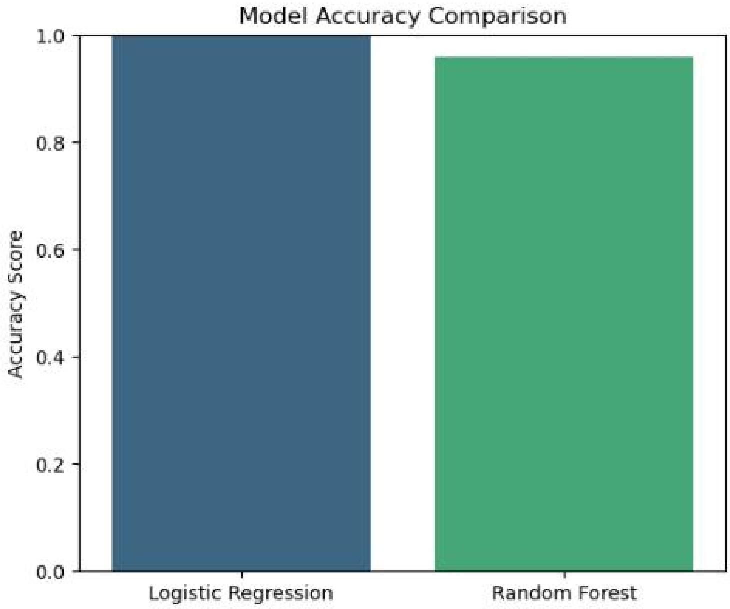


The Random Forest confusion matrix shows a few misclassifications in classes 1 and 2, but overall classification is strong. Class 1 had slight underestimation, and class 2 had a small number of false negatives, which accounts for the slightly lower overall accuracy.

5.4. Model Accuracy Comparison

The bar chart compares the accuracy of Logistic Regression and Random Forest. Logistic Regression reaches perfect accuracy, while Random Forest achieves slightly lower performance but remains robust. This visualization emphasizes the comparative strengths of both models.

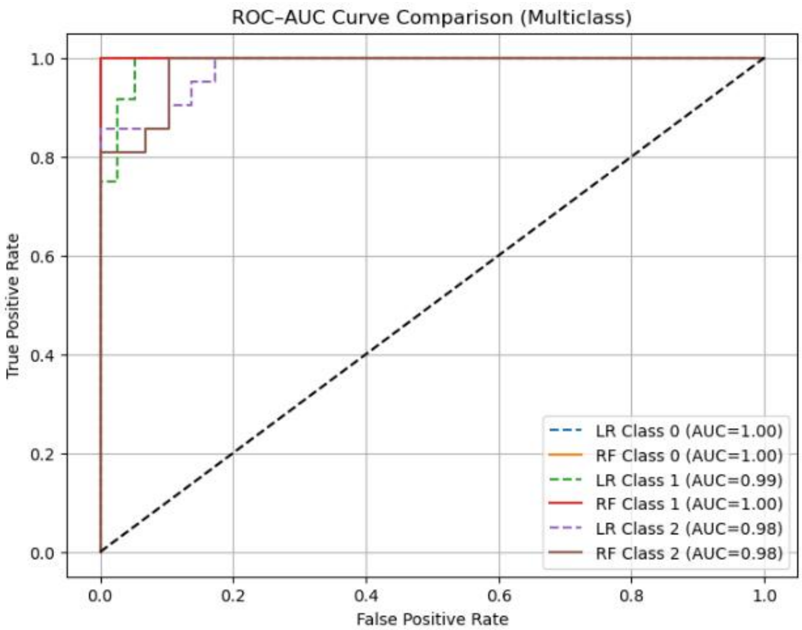
Figure 3. Model Accuracy Comparison



5.5. Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC) Curves

ROC-AUC curves for both models illustrate the sensitivity-specificity trade-off for each class. Logistic Regression shows AUC of 1.0 for all classes, indicating perfect discrimination. Random Forest demonstrates slightly lower AUC for class 2, reflecting minor misclassification and confirming the numerical accuracy metrics.

Figure 4. ROC-AUC Curves Comparison (Multiclass)



5.6 Discussion

The results indicate that both Logistic Regression and Random Forest are effective for predicting breast cancer risk. Logistic Regression's perfect accuracy and AUC values suggest strong linear separability within the dataset. Random Forest captures complex feature interactions, providing robustness in real-world scenarios, although minor misclassifications are observed. Integrating predictions from both models can further enhance early detection and risk stratification. Confusion matrices, accuracy comparisons, and ROC-AUC curves collectively support the use of these models in clinical decision-making, timely medical referral, and preventive healthcare strategies.

6.0 Conclusions

This study highlights the importance of Breast Self-Examination (BSE) as a simple and effective method for early detection of breast abnormalities. By combining socio-economic and clinical data with stochastic and machine learning techniques, the research successfully developed predictive models that classify breast cancer risk with high accuracy. Logistic Regression achieved perfect accuracy, showing that the dataset followed a clear linear pattern, while Random Forest demonstrated strong performance in handling complex relationships between features. These results confirm that mathematical and stochastic models can provide valuable insights for preventive health analysis without the need for advanced artificial intelligence systems. The integration of BSE awareness, clinical examination, and computational modeling can support healthcare professionals in identifying high-risk individuals at an early stage. Overall, the study emphasizes that public health education, regular self-examination, and simple predictive tools can together reduce the burden of breast cancer and enhance women’s health outcomes in India and beyond.

References

1. Kandasamy, G., Almaghaslah, D., Almanasef, M., & Alamri, R. D. A. (2024). Knowledge, attitude, and practice towards breast self-examination among women: a web-based community study. Frontiers in Public Health, 12, 1450082.
2. Yildirim, H., & Yildirim, S. (2022). A different approach to breast self-examination training. Global Medical Journal, 14(1), E202236. https://ifnmujournal.com/gmj/article/view/E202236
3. Francks, L., Murray, A., & Wilson, E. (2023). Barriers and facilitators to breast self-examination in women under 50 in an international context: A qualitative systematic review. International Journal of Health Promotion and Education, 1–18.
4. Dechasa, D. B., Asfaw, H., & Abdisa, L. (2022). Practice of breast self-examination and associated factors among female health professionals working in public hospitals of Harari Regional State, Eastern Ethiopia. Frontiers in Oncology, 12, 1002111.
5. Jalloul, R. (2023). A Review of Machine Learning Techniques for the Detection of Breast Cancer. Journal of Cancer Research & Clinical Oncology.
6. Li, C. (2023). A Systematic Review of Application Progress on Machine Learning Models in Breast Cancer Diagnosis. Journal of Cancer Research & Clinical Oncology.
7. Rabiei, R. (2022). Prediction of Breast Cancer using Machine Learning Approaches. Journal of Cancer Research & Clinical Oncology.
8. Jafari, A. (2024). Machine-learning methods in detecting breast cancer and defining its type. Journal of Cancer Research & Clinical Oncology.
9. Islam, T. (2024). Predictive modeling for breast cancer classification in the Bangladeshi population. Scientific Reports.
10. Chen, T. (2024). Using an innovative method for breast cancer diagnosis. Journal of Cancer Research & Clinical Oncology.
11. Lee, S. (2006). A stochastic model for predicting the mortality of breast cancer. Journal of the National Cancer Institute Monographs.
12. Martinez, R. G. (2023). Deep learning algorithms for the early detection of breast cancer. Journal of Cancer Research & Clinical Oncology. https://www.sciencedirect.com/science/article/pii/S2352914823001636
13. Siah, K. W. (2019). Machine-learning and stochastic tumor growth models for predicting breast cancer outcomes. Clinical Cancer Investigation Journal.
14. Rahman, M. A. (2025). Advancements in Breast Cancer Detection: A Review of Global Trends, Risk Factors, Imaging Modalities, Machine Learning, and Deep Learning Approaches. BioMedInformatics.

15. Ahmed, K. A. (2025). Advancing breast cancer prediction: Comparative analysis of machine learning algorithms. PLOS ONE.
16. Gurmessa, D. K. (2024). Explainable machine learning for breast cancer diagnosis from mammography and ultrasound images. BMJ Health & Care Informatics.
17. Ghasemi, A. (2024). Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review.