Explainable AI-Based Breast Cancer Prediction and Decision Support System

K. Sathvika Reddy
UG Student, Dept of CSE (Data Science) Vidya Jyothi Institute of Technology Hyderabad, Telangana, India ksathvikareddy4@gmail.com

M Retwwik Krishna Yadav
UG Student, Dept of CSE (Data Science) Vidya Jyothi Institute of Technology Hyderabad, Telangana, India ritvikyadav213213@gmail.com

K. Trivikram Reddy
UG Student, Dept of CSE (Data Science) Vidya Jyothi Institute of Technology Hyderabad, Telangana, India trivikramreddy868@gmail.com

M. Prasanna Kumari
Assistant Professor Dept of CSE (Data Science) Vidya Jyothi Institute of Technology Hyderabad, Telangana, India prasannamcse@vjit.ac.in

K. Nakshathra Reddy
UG Student, Dept of CSE (Data Science) Vidya Jyothi Institute of Technology Hyderabad, Telangana, India nakshathrareddy4@gmail.com

How to Cite this Article:

Reddy, K. S., Yadav, M. R. K., Reddy, K. T. & Reddy, K. N. (2026). Explainable AI-Based Breast Cancer Prediction and Decision Support System. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(04). https://doi.org/10.55041/ijcope.v2i4.943

https://doi.org/10.55041/ijcope.v2i4.943

ABSTRACT— Breast cancer is one of the leading causes of death among women worldwide, and early detection is essential for improving survival rates. This project presents an Explainable Artificial Intelligence (XAI)-based Breast Cancer Prediction and Decision Support System using the Breast Cancer Wisconsin dataset, where machine learning algorithms such as Logistic Regression, Decision Trees, and Random Forest are applied to classify tumors as benign or malignant based on features extracted from fine needle aspirate images. To overcome the limitations of traditional black-box models, the system integrates explainability techniques like SHAP and LIME, which provide clear insights into how different features influence the model's predictions, thereby enhancing transparency and trust. The model is trained and tested on the dataset to achieve high accuracy while also offering interpretable results, enabling healthcare professionals to understand the reasoning behind each prediction. Overall, the proposed system not only improves diagnostic accuracy but also supports clinical decision-making by combining reliable predictions with meaningful explanations, making it a valuable tool for early breast cancer detection and treatment planning.

INTRODUCTION

Breast cancer is one of the most prevalent and life- threatening diseases affecting women across the globe, accounting for a significant number of cancer-related deaths each year. According to global health studies, early detection and timely treatment can drastically improve survival rates and reduce mortality. However, accurate diagnosis often depends on the expertise of medical professionals and the availability of advanced diagnostic tools. In recent years, the rapid growth of Artificial Intelligence (AI) and MachineLearning (ML) has opened new possibilities in the field of healthcare, particularly in disease prediction and diagnosis. These technologies enable the analysis of large volumes of medical data to identify patterns and make predictions with high accuracy, thereby assisting doctors in making informed decisions.

One of the most widely used datasets for breast cancer prediction is the Breast Cancer Wisconsin dataset, which contains features computed from digitized images of fine needle aspirate (FNA) of breast masses. These features describe characteristics such as radius, texture, perimeter, area, smoothness, and concavity of the cell nuclei, which are crucial indicators in distinguishing between benign and malignant tumors. Machine learning models trained on this dataset have shown promising results in classifying tumors, thus supporting early diagnosis and reducing the chances of human error.

Despite the high accuracy of many ML models, a major challenge remains: most of these models function as "black boxes," meaning they provide predictions without explaining how the decisions are made. In critical domains like healthcare, this lack of transparency can limit trust and hinder adoption by medical professionals. Doctors need to understand the reasoning behind a prediction before relying on it for clinical decisions. This is where Explainable Artificial Intelligence (XAI) plays a vital role. XAI focuses on making AI models more interpretable by providing clear explanations of how input features contribute to the output.

In this project, an Explainable AI-based Breast Cancer Prediction and Decision Support System is developed using the Breast Cancer Wisconsin dataset. The system employs machine learning algorithms such as Logistic Regression, Decision Trees, and Random Forest to classify tumors as benign or malignant. Additionally, it integrates explainability techniques like SHAP (Shapley Additive Explanations) to provide insights into model predictions. These techniques help identify the most influential features and explain individual predictions in a human-understandable manner.

I. PROBLEM DEFINITION

Breast cancer diagnosis requires high precision and reliability, as incorrect or delayed predictions can have serious consequences on patient health. Although machine learning models have shown high accuracy in predicting breast cancer, their lack of interpretability makes them difficult to trust and adopt in real-world clinical settings. Most existing systems function as black-box models, offering little to no explanation for their predictions, which limits their usefulness for healthcare professionals.

Therefore, the problem is to design and develop a breast cancer prediction system that not only provides accurate classification of tumors as benign or malignant but also ensures transparency by explaining the reasoning behind each prediction. The system should support doctors in clinical decision-making by identifying key influencing factors and presenting results in an understandable manner.

1.2 OBJECTIVES

The main objectives of this project are as follows:

- To develop a machine learning-based model for predicting breast cancer using the Breast Cancer Wisconsin dataset.
- To accurately classify tumors as benign or malignant based on extracted features.
- To integrate Explainable AI techniques such as SHAP and LIME to provide clear and interpretable explanations for model predictions.
- To identify and analyze the most important features influencing the classification results.
- To design a decision support system that assists healthcare professionals in understanding and

trusting AI-based predictions.

- To improve the overall transparency, reliability, and usability of AI in medical diagnosis.

1.3 LITERATURE SURVEY

Artificial Intelligence (AI) and Machine Learning (ML) have greatly improved healthcare, especially in early breast cancer detection. Traditional methods are time-consuming and depend on expert analysis, so researchers use models like Logistic Regression, Decision Trees, SVM, and Random Forest with the Breast Cancer Wisconsin dataset. These models achieve very high accuracy (up to ~99%) due to well- structured data. The dataset, with 569 samples and 30 features, is widely used because it is clean, reliable, and suitable for both basic and advanced models.

Deep learning techniques such as Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN) further improve prediction by capturing complex patterns in data. However, both machine learning and deep learning models often act as "black-box" systems, meaning their decisions are not easily understood. This lack of interpretability reduces trust among doctors and limits their use in real clinical settings, creating the need for Explainable Artificial Intelligence (XAI).

To solve this, researchers combine prediction models with explainability methods like SHAP and LIME, which show how features influence results. Interpretable models such as Generalized Additive Models (GAM), Explainable Boosting Machines (EBM), and Decision Trees are also used. These hybrid and explainable approaches improve transparency, support better medical decisions, reduce errors, and help identify important disease factors, making AI more practical and trustworthy in healthcare.

II. METHODOLOGY

1. Data Collection
Breast Cancer Wisconsin dataset is used, containing features of cell nuclei to classify tumors as benign or malignant.

2. Data Preprocessing
Data is cleaned, missing values are handled, features are normalized using MinMax scaling, and the dataset is split into training and testing sets.

3. Feature Selection
Important features are selected using methods like correlation analysis and statistical techniques, and irrelevant features are removed to improve model performance.

4. Model Training
Machine learning models such as Random Forest, Support Vector Machine (SVM), or Logistic Regression are trained to learn patterns in the dataset.

5. Explainability Integration
Explainable AI techniques like SHAP are applied to understand feature contributions and provide transparency in predictions.

6. Classification & Evaluation
The model classifies tumors as benign or malignant and is evaluated using accuracy, precision, recall, F1-score, and confusion matrix.

III. PROPOSED SYSTEM

The proposed system is an **Explainable AI-based Breast Cancer Detection and Decision Support System** designed to accurately classify tumors as benign or malignant while providing clear explanations for its predictions. The system uses the Breast Cancer Wisconsin dataset as input and applies advanced machine learning techniques along with explainability methods to improve both performance and transparency.

The system begins with data preprocessing, where the dataset is cleaned, normalized, and prepared for analysis. Important features are selected to improve efficiency and reduce complexity. A machine learning model, such as Random Forest or Support Vector Machine (SVM), is then trained to identify patterns in the data and make accurate predictions.

.

IV. IMPLEMENTATION DETAILS

The implementation of the proposed intrusion The proposed system is implemented using Python with libraries such as NumPy, Pandas, Scikit-learn, Matplotlib, and SHAP. The Breast Cancer Wisconsin dataset is loaded and processed using Pandas, where data preprocessing is performed by handling missing values, normalizing features using MinMaxScaler, and splitting the dataset into training and testing sets. Machine learning models such as Random Forest, Support Vector Machine (SVM), or Logistic Regression are then trained on the dataset, with techniques like hyperparameter tuning applied to improve performance. To enhance interpretability, SHAP (Shapley Additive Explanations) is integrated to analyze feature contributions and provide visual explanations of predictions. Finally, the system evaluates performance using metrics such as accuracy, precision, recall, F1-score, and confusion matrix, and produces an output that classifies tumors as benign or malignant along with an explanation to support clinical decision-making.
ALGORITHMS USED

4.1.1 LOGISTIC REGRESSION

Logistic Regression is a supervised learning algorithm used for binary classification. It uses a sigmoid function to convert outputs into probabilities between 0 and 1. It serves as a baseline model due to its simplicity and interpretability. It works well with linearly separable data but is not suitable for complex nonlinear relationships.

4.1.2 DECISION TREE

A Decision Tree is a tree-based algorithm that splits data based on feature values to make decisions. It provides rule-based classification and is easy to understand and visualize. However, it is prone to overfitting, especially with large datasets.

4.1.3 SUPPORT VECTOR MACHINE (SVM)

SVM is a powerful algorithm used for both linear and nonlinear classification. It works by finding an optimal hyperplane that separates data points. It performs well in high-dimensional spaces but is computationally expensive and difficult to interpret.

4.1.4 RANDOM FOREST (RF)

Random Forest is an ensemble method that combines multiple decision trees and uses voting for final prediction. It improves accuracy and reduces overfitting. It is robust and handles large datasets effectively but is less interpretable than a single decision tree.

4.1.5 SHAP (Shapley Additive Explanations)

SHAP is an Explainable AI technique based on game theory that assigns importance values to each feature. It helps explain both global and individual predictions with high accuracy but requires high computational power.

4.1.6 LIME (Local Interpretable Model-Agnostic Explanations)

LIME is an XAI method that explains individual predictions by approximating the model locally. It is simple and model- agnostic but only provides local explanations and not a complete global understanding.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The Explainable AI-Based Breast Cancer Prediction System produces both predictive results and interpretable explanations, making it useful for clinical decision support.

5.1 Prediction Output

The primary output of the system is the classification of the tumor based on input features:

- Benign (0): Non-cancerous tumor
- Malignant (1): Cancerous tumor

The system also provides a probability score, indicating the confidence level of the prediction (e.g., 95% malignant).

5.2 Performance results

The machine learning models are evaluated using standard metrics:

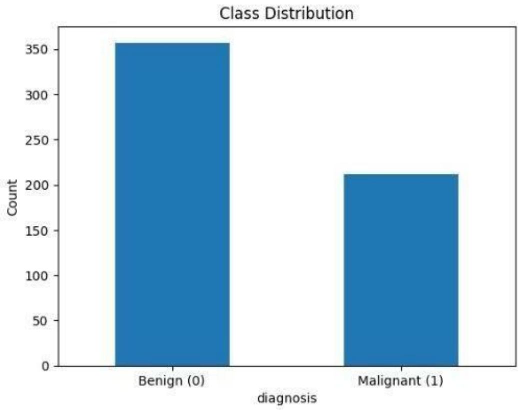
- **Accuracy:** Measures overall correctness (typically above 95%)
- **Precision:** Indicates the correctness of positive predictions
- **Recall (Sensitivity):** Measures the ability to detect malignant cases
- **F1-Score:** Balances precision and recall

Among the models used, Random Forest and SVM generally provide the highest accuracy and reliability for this dataset.

```
Accuracy: 0.9649122807017544

Confusion Matrix:
[[70  1]
 [ 3 40]]

Classification Report:
              precision    recall  f1-score   support

           0       0.96      0.99      0.97        71
           1       0.98      0.93      0.95        43

    accuracy                           0.96       114
   macro avg       0.97      0.96      0.96       114
weighted avg       0.97      0.96      0.96       114
```

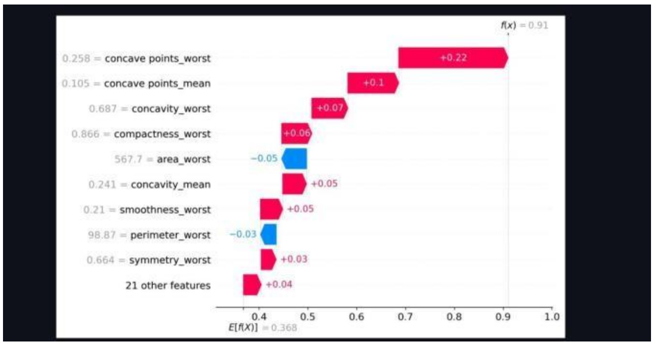
Classification Report

Fig. 1. Accuracy Graph



Class Distribution of Tumor Diagnosis (Benign vs Malignant)

Fig, 2. SHAP Analysis for Model Performance



SHAP Analysis for Model Performance

VI. CONCLUSION

Breast cancer remains a major global health challenge where early detection and accurate diagnosis are crucial for improving patient survival rates. This project presents an Explainable AI-based Breast Cancer Prediction and Decision Support System using the Breast Cancer Wisconsin dataset, where machine learning models such as Logistic Regression, Decision Trees, and Random Forest are used to classify tumors as benign or malignant with high accuracy, thereby supporting early detection and reducing diagnostic errors. A key contribution of this work is the integration of Explainable Artificial Intelligence (XAI) techniques like SHAP and LIME, which provide clear explanations of how features influence predictions, addressing the limitations of black-box models and enhancing transparency, trust, and clinical acceptance. The system not only improves prediction performance but also supports better decision-making by identifying important features and enabling personalized treatment planning. Additionally, it can assist in preliminary diagnosis in resource- limited settings and improve communication between doctors and patients. However, the system has limitations such as reliance on a single dataset and lack of real-time clinical validation, which can be addressed in future work by incorporating diverse datasets, real-world data, and advanced deep learning models, along with deployment as a web or mobile application. Overall, the proposed system effectively combines accuracy and interpretability, making it a reliable, transparent, and practical tool for real-world healthcare applications.

VII. FUTURE SCOPE

The proposed system can be further enhanced by incorporating larger and more diverse datasets from multiple medical sources to improve accuracy and generalization. Future work can focus on integrating advanced deep learning models such as Convolutional Neural Networks (CNN) and hybrid architectures to capture more complex patterns in medical data. The system can also be extended for real-time implementation by developing a web or mobile-based application that can be integrated with hospital information systems for practical clinical use. Additionally, incorporating other Explainable AI techniques and improving visualization methods can provide deeper insights into model predictions. The system can be expanded to support multi-class classification and detect different stages of breast cancer, enabling more detailed diagnosis. Further improvements may include integrating medical imaging data such as mammograms and histopathological images to enhance prediction capability. Overall, these advancements can make the system more robust, scalable, and suitable for real-world healthcare applications..

VIII. ACKNOWLEDGMENT

We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project, we take this opportunity to express our profound gratitude and deep regard to our guide **Mrs. M. PRASANNA KUMARI** Designation for his/her exemplary guidance, monitoring and constant encouragement throughout the project work. The blessing, help and guidance given by him/her shall carry us a long way in the journey of life on which we are about to embark.

We are also thankful to **Dr. K.S.R.K. SARMA** Head, Department of Computer Science and Engineering (data science) for providing encouragement and support for completing this project successfully.

Finally, we would like to take this opportunity to thank our family for their constant encouragement, without which this assignment would not be completed. We sincerely acknowledge and thank all those who gave support directly and indirectly in the completion of this project.

IX. REFERENCES

[1] Tarek Khater; Abir Hussain; Riyad Bendardaf; Iman M. Talaat; Hissam Tawfik; Sam Ansari(2023)
An Explainable Artificial Intelligence Model for the Classification of Breast Cancer
Available at: https://ieeexplore.ieee.org/abstract/docum ent/10229149
[2] Shalini Raghuvansi; Yashs Tiwari; Abhishek Yadav(2025)
Decoding Breast Cancer Detection: An Explainable AI Framework for Transparent Diagnosis
Available at:https://ieeexplore.ieee.org/abstract/document/11088047

[3] Zulfikar Ali Ansari, Manish Madhava Tripathi & Rafeeq Ahmed(2025)
The role of explainable AI in enhancing breast cancer diagnosis using machine learning and deep learning models
Available at: https://link.springer.com/article/10.1007/s44163-025- 00307-8

[4] Vinayak Gupta; Richa Sharma(2024)

Enhancing Breast Cancer Prediction with XAI-Enabled Boosting Algorithms

Available at: https://ieeexplore.ieee.org/abstract/document/10723843
[5] Olya Rezaeian, Ehsannodin Ghorbanichemazkati, A. Emrah Bayrak & Onur Asan(2026)

Clinician perspectives on trust and adoption of AI in breast cancer diagnosis
Available at: https://www.tandfonline.com/doi/abs/10.1080/24725579.20 25.2611435

[6] Mubasshira, Md Mahbubur Rahman, Lisa Akter (2026) **Artificial Intelligence (AI)-Assisted Treatment of Breast Cancer**
Available at:
https://link.springer.com/chapter/10.1007/978-981-95-3682- 5_18

[7] Inam Ul Haq, Akib Mohi Ud Din Khanday(2025) **Exploring the Role of Explainable AI in Women's Health**
Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/978139424 9312.ch13

[8] Abdurakhimova L. A., Komilova L. B.(2026) **ARTIFICIAL INTELLIGENCE-DRIVEN ADVANCES IN BREAST CANCER DETECTIONS,DIAGNOSIS AND PERSONALIZED TREATMENT**
Available at: http://eurekaoa.com/index.php/5/article/view/244